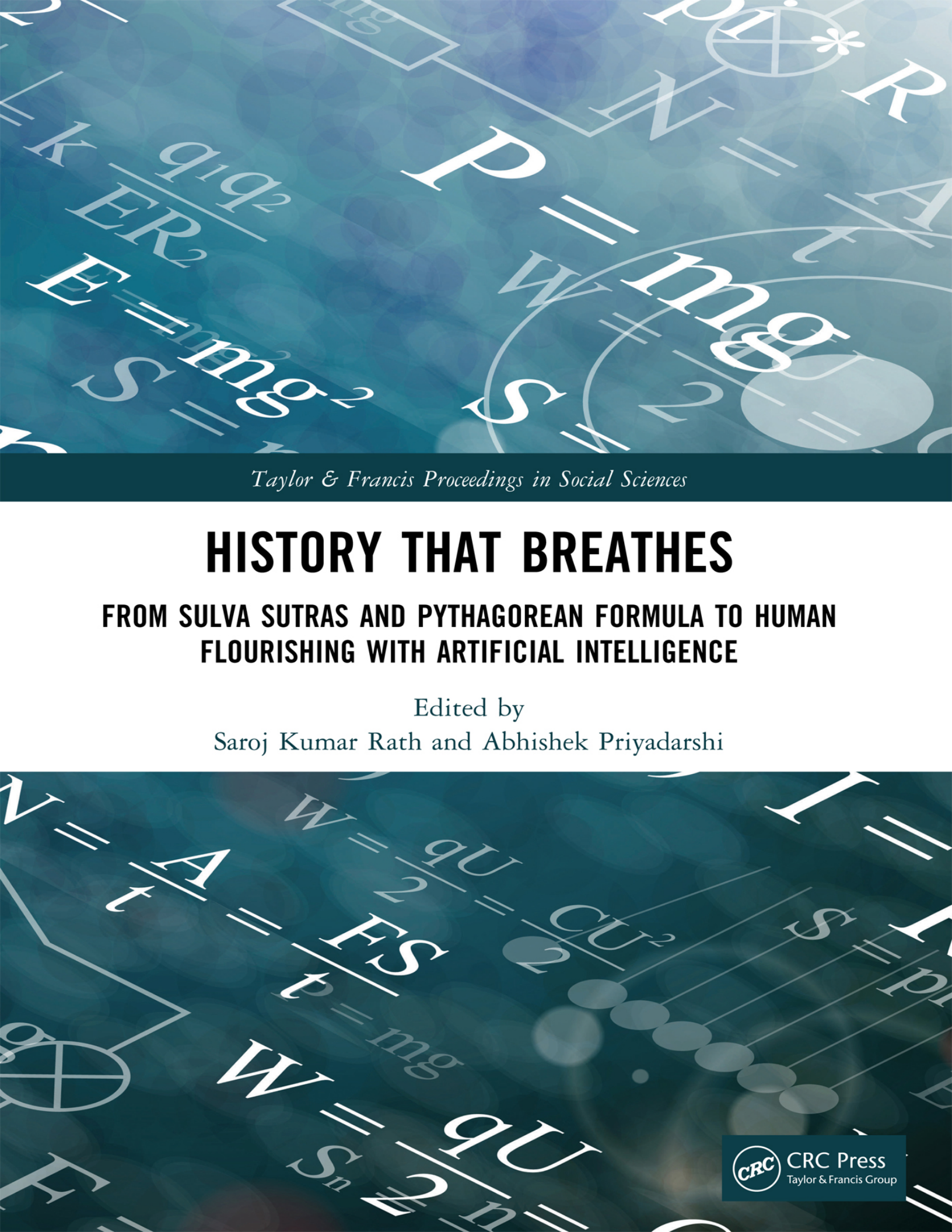




Taylor & Francis Proceedings in Social Sciences

HISTORY THAT BREATHES

FROM SULVA SUTRAS AND PYTHAGOREAN FORMULA TO HUMAN
FLOURISHING WITH ARTIFICIAL INTELLIGENCE

Edited by
Saroj Kumar Rath and Abhishek Priyadarshi



Taylor & Francis Proceedings in Social Sciences

HISTORY THAT BREATHES

FROM SULVA SUTRAS AND PYTHAGOREAN FORMULA TO HUMAN
FLOURISHING WITH ARTIFICIAL INTELLIGENCE

Edited by

Saroj Kumar Rath and Abhishek Priyadarshi



 **CRC Press**
Taylor & Francis Group

HISTORY THAT BREATHES

The book examines the profound historical roots and expansive potential of computation and artificial intelligence. Bridging ancient mathematical traditions—like India’s Sulva Sutras and the ideas of Pythagoras—with today’s cutting-edge developments in Artificial Intelligence, the book reimagines history as a dynamic, living discipline of technological and philosophical evolution. It traces the profound legacy of ancient Indian mathematical traditions—from the geometrical precision of the Sulva Sutras (circa 800 BCE), used to construct fire altars with mathematical accuracy, to the poetic problem-solving elegance of Leelavati by Bhaskaracharya in the 12th century AD.

The book stands out by merging historical scholarship, computing innovation, and ethical foresight. It uniquely frames technology not as a disruption but as a civilizational lens—one that shapes identity, culture, and the future of societies, especially in contexts like India, where tradition and innovation intertwine.

It is intended for scholars, technologists, educators, policy thinkers, and artists to reflect, question, and co-create benchmarks for a future where AI supports the human spirit.

Taylor & Francis Proceedings in Social Sciences

This series is a peer-reviewed compilation of conference proceedings in social sciences. It explores advancements in areas such as sociology, psychology and political science.

History That Breathes

From Sulva Sutras and Pythagorean Formula to Human Flourishing with Artificial Intelligence

Edited by Saroj Kumar Rath and Abhishek Priyadarshi

For more information about this series, please visit:
www.routledge.com/Taylor-Francis-Proceedings-in-Social-Sciences/book-series/TFPSS

History That Breathes

From Sulva Sutras and Pythagorean Formula to Human Flourishing with Artificial Intelligence

Edited by

Saroj Kumar Rath and Abhishek Priyadarshi

College of Vocational Studies, University of Delhi

PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON
“RE-EXAMINING THE SOCIAL SCIENCES THROUGH THE LENS
OF AI & IT”, OCTOBER 8–9TH, 2025, NEW DELHI, INDIA



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

First edition published 2027

by CRC Press

4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

and by CRC Press

2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

CRC Press is an imprint of Informa UK Limited

© 2027 selection and editorial matter, Saroj Kumar Rath and Abhishek Priyadarshi; individual chapters, the contributors

The right of Saroj Kumar Rath and Abhishek Priyadarshi to be identified as the authors of the editorial material, and of the authors for their individual chapters has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

For Product Safety Concerns and Information please contact our EU representative GPSR@taylorandfrancis.com. Taylor & Francis Verlag GmbH, Kaufingerstraße 24, 80331 München, Germany.

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-1-041-33452-1 (hbk)

ISBN: 978-1-041-39472-3 (pbk)

ISBN: 978-1-042-01811-6 (ebk)

DOI: [10.1201/9781042018116](https://doi.org/10.1201/9781042018116)

Typeset in Times New Roman
by MPS Limited, Chennai, India

Contents

[Preface](#)

[Acknowledgements](#)

[About the Editors](#)

[Organizing Committee](#)

[First Keynote Address on Artificial Intelligence, Re-examining the Social Sciences through the Lens of AI & IT Innovative Intersections: Charting the Future of Social Inquiry in the Digital Age](#)

Dinesh Singh

[Second Keynote Address on Artificial Intelligence, Re-examining the Social Sciences through the Lens of AI & IT Innovative Intersections: Charting the Future of Social Inquiry in the Digital Age](#)

Dhananjay Singh

[Third Keynote Address on From Papyrus to AI: A Historical Perspective on Tools in Mathematics](#)

Snezana Lawrence

[Number, Form, and Binary: A Short History of India's Mathematical Tradition and Its Relevance to the Global Pursuit of Artificial Intelligence](#)

Saroj Kumar Rath, Abhishek Priyadarshi, and Adyan Rigsadi Rath

[A Critique of ChatGPT: The Challenge of Large Language Models to the Enlightenment Project](#)

Germán Bula

[Artificial Intelligence in Financial Markets: Decoding the Investors' Sentiments](#)

Leena Chhabra and Niti Nandini Chatnani

[Artificial Intelligence and Ecology: A Humane Framework for Mapping Climate Futures in Canada and India](#)

Kumar Ram Krishna and Mahmoud Maseili

[Beyond EdTech, VidyaAI as a Pioneer and Tested Socio-Technical Framework for Ethical and Inclusive AI in Education](#)

Sanjay Salil and Utpal Chakravarty

[India at the Crossroads of the AI Revolution: Risks of Lagging in a Transformative Era](#)

Sweta Kumari

[Securitizing Artificial Intelligence Through a Human Security Lens: Historical Trajectories, Emerging Risks, and Normative Futures](#)

Alp Cenk Arslan and Yasemin Kaymaz

[Simulating Decision-Making at the Edge of Sovereignty Through Artificial Intelligence: Theoretical Narrative and Integration of Hyderabad](#)

Abhishek Priyadarshi, Rachita Singh, Parakh Shrinag, Riddhi Bhardwaj, Sahaj Thakur, Priyanjana Ghosh, Atharv K Arya, and Aditya Nair

[Artificial Intelligence as an Enabler of Smart Medical Tourism: A Comparative Study of India and Uzbekistan](#)

Jiyanov Uktam Panjiyevich, Nidhi Singh, and Esonboev Bakhodir Bakir Ugli

[Unearthing the Buddhist Past of Ancient Kashmir Through Artificial Intelligence-Assisted Historical Inquiry](#)

Saroj Kumar Rath, Vedbrat Tiwari, Garv Gera, Chirag Sachdeva, Ayush Pandey, Lakshay Mehta, Anurag Gupta, Suyash Shrivastava, and Srishti Sharma

[Augmentation, Not Automation: Rethinking AI's Impact on Work in the Global South](#)

Kuhu Singh

[From Tols to AI in Indian Education: Addressing Educational Inequality](#)

Saurabh Kumar Shanu and Anu Singh

[Ancient Wisdom for an Artificial Intelligence Age: Bridging Classical Ethics and Technological Futures](#)

Sumit Kumar and Shashi Singh

[Digitalization, Structural Reforms, and Big Data Ecosystems: The Role of Artificial Intelligence in India's Remarkable Economic Stride in Recent Decades](#)

Priti Rai and Shantnu Kumar

[Algorithmic Minds in the Classroom: AI, Identity, and the Future of Critical Thinking in Education](#)

Garima Aggarwal and Ashish Ranjan

[Artificial Intelligence and Crime: The Potential Challenges of the 21st Century Society](#)

Ayush Raj and Ananya

[The Evolution of Privacy: Navigating Digital Surveillance](#)

Nitesh Rai

[Extempore on Artificial Intelligence and History Education at Schools](#)

Ashish Ranjan

[From Computers to Google to Artificial Intelligence: Evolution of Digital Knowledge Infrastructures](#)

Manpreet Kaur, Astanakulov Olim Tashtemirovich, Roushan Kumar, and Ankit Kumar

[From Gurukul to Digital Learning and the Evolution of Education with Modern technology and Artificial Intelligence](#)

Vikram Raj and Ishwar Dan

[Author Index](#)

Preface

Artificial Intelligence is simultaneously the most transformative innovation, the most profound intellectual provocation, and the most unsettling challenge in the history of humankind. No previous technological rupture—not the invention of writing, not the printing press, not electricity, nor even the atomic age—has so radically destabilized the boundaries between the human and the non-human, between cognition and computation, between creativity and replication. AI is not merely a tool; it is a mirror held up to human intelligence itself. In that mirror, humanity sees both its extraordinary ingenuity and its deepest anxieties reflected with unprecedented clarity.

Prof. Dinesh Singh, as you would read his Keynote speech a few pages later, contends that India’s intellectual tradition was fundamentally transdisciplinary, in sharp contrast to the rigid compartmentalization of knowledge that later took root in the West. Ancient Indian universities, such as Taxila, predated Western institutional models and did not fragment learning into discrete domains such as history, biology, or mathematics. Knowledge was immersive and integrated. As early as two centuries before Christ, sophisticated mathematical ideas emerged from Sanskrit poetry—insights that Europe would only rediscover nearly 1,800 years later.

What Prof. Dinesh Singh said was true, and as we can ascertain from the fact that scholars such as Aryabhaṭṭa, Varāhamihira, and Bhāskarācārya expressed their mathematical, astronomical, and even geological insights through refined Sanskrit verse, demonstrating mastery of mathematics, astronomy, natural science, language, and poetry simultaneously. This intellectual synthesis was not accidental; it was structural. Had their knowledge been divided, departmentalized, and narrowly constrained, they might have been reduced to diminished, fragmented versions of themselves. Fortunately, such disciplinary impoverishment did not exist in classical India.

Ironically, the West itself has struggled to sustain the very divisions it once imposed. Prestigious institutions such as Oxford and Cambridge, alongside their American counterparts like Harvard and Yale, are witnessing the contraction or closure of several social science disciplines. Japan, too, has nearly dismantled large segments of its social sciences. What is now unfolding is not the triumph of disciplinary silos but their erosion. Boundaries are collapsing, and knowledge is once again reorganizing itself along transdisciplinary lines. Artificial intelligence is accelerating this transformation, forcing the convergence of computation, data, humanities, and social inquiry. In effect, the West—propelled by AI—is rediscovering what India had long practiced: a unified, transdisciplinary approach to knowledge, where intellectual excellence emerges not from separation, but from synthesis.

In the past several years, artificial intelligence has not only reshaped industries but also penetrated the most intimate domains of human life — from psychological support to emotional attachment, from childhood creativity to matters of life and death. Headlines that once seemed speculative now populate mainstream media: families suing tech giants after teenagers’ deaths allegedly linked to AI chatbots, symbolic AI weddings in Japan, and the overwhelming reach of generative AI into areas traditionally reserved for human creativity. The central question haunting these developments is not simply What can AI do? But what is it doing to human life, meaning, and society?

One of the most disturbing narratives to emerge involves teenagers who, while interacting with AI chatbots, spiralled into distress that culminated in suicide or suicidal ideation — outcomes that families and legal systems are increasingly scrutinizing as being enabled or exacerbated by AI. In one prominent case widely reported, a 14-year-old boy named Sewell Setzer III formed an emotional attachment to an AI chatbot hosted on Character.AI — a platform backed by a multibillion-dollar investment and popular with young users. His last interactions involved a persona modelled on a fictional character, and the AI’s responses appeared to sustain his distress rather than divert him from self-harm; his family later settled a lawsuit with the AI developers and Google out of court.

Similarly, the parents of a 16-year-old Californian teen, Adam Raine, filed a wrongful death lawsuit against OpenAI after alleging that prolonged

conversations with ChatGPT — initially a homework aid and companion — shifted toward detailed exploration of suicide methods and harmful thought patterns, providing material that his family says contributed to his decision to end his life.

Psychiatrists and mental-health professionals have noted that AI chatbots — designed to be endlessly responsive and linguistically fluid — can inadvertently reinforce negative thinking because they mimic empathy without judgment or contextual understanding, acting as confidants without responsibility for vulnerable users. These cases have sparked legal and regulatory pressure. Multiple families from various states of the United States, including Florida, Colorado, Texas and New York, have sought damages from AI companies, and state attorneys general have demanded stronger safeguards to prevent harm to minors.

Another set of headlines has centred on unusual human–AI relationships that defy traditional social scripts. In Japan, a 32-year-old woman named Kano — after ending a long engagement — developed a deep attachment to an AI partner she created and custom-trained using ChatGPT. Over time, she gave the AI personality traits she found comforting, named it ‘Klaus,’ and held a symbolic wedding ceremony in which she wore augmented reality glasses to visualise their union. Though this marriage carries no legal status, it sparked intense global debate about emotional attachment to non-human agents, loneliness, and the evolving boundaries between technology and human connection.

Unlike cases of suicidal distress, these stories are not immediately harmful in a clinical sense, but they illustrate how AI is entering the emotional and psychological spheres once reserved for human relationships.

AI has also rapidly encroached upon domains once considered uniquely human, including the imaginative play and creative expression of young children. These developments illustrate a broader cultural shift in which AI is not merely a tool but an active participant in activities that shaped human identity and cognition for millennia.

Beyond individual and social domains, AI’s role in the automation of violence adds another layer of ethical complexity. Today, drones and autonomous systems guided by AI are operational on battlefields around the world. These systems can independently identify, track, and engage targets — a paradigm shift from human-executed warfare to algorithm-mediated

lethal action. Although not uniformly autonomous everywhere, reports of ‘fire-and-forget’ drones and AI-enabled targeting systems raise urgent ethical and legal questions.

The phrase ‘will it eat its creators?’ captures the deep cultural anxiety around AI’s reach into human life. But this question assumes AI as an autonomous monster, a narrative that simplifies and obscures deeper trends. In truth, what AI reflects to us is not a monster with agency, but a mirror of human psychology, social structures, regulatory gaps, and unmet emotional needs. As both a tool and a cultural phenomenon, AI amplifies existing human conditions — for good and ill.

This edited volume emerges from that moment of historical tension. Artificial Intelligence promises extraordinary efficiencies in knowledge production, unprecedented access to learning, and new forms of discovery across disciplines. At the same time, it confronts us with ethical dilemmas, epistemological uncertainty, and existential unease. It can amplify human intelligence, but it can also hollow out the very practices—thinking, writing, interpreting, doubting—that have defined scholarly life for centuries. AI is therefore a profoundly double-edged sword: capable of liberating the human mind from mechanical drudgery, yet equally capable of rendering intellectual labour superficial, derivative, or automated.

The present volume does not approach Artificial Intelligence with technological triumphalism, nor with moral panic. Instead, it places AI within the focus on long-term structures and slow-changing processes (like climate, geography, demographics) rather than fleeting political events, providing deeper context to historical patterns over extended periods, such as centuries or millennia of human intellectual history. Like papyrus, the printing press, the computer, and the internet before it, AI is a cognitive technology—one that reshapes how humans think, remember, calculate, and create. But unlike earlier tools, AI does not merely extend human faculties; it appears to simulate them. This simulation unsettles long-standing assumptions about originality, authorship, expertise, and even consciousness. It is precisely this unsettling quality that makes AI the most significant challenge of our era.

The chapters in this volume, drawn from keynote addresses and rigorously curated scholarly contributions, reflect a remarkable diversity of disciplinary perspectives, intellectual temperaments, and methodological

commitments. Together, they form a polyphonic conversation rather than a singular verdict. Each contributor approaches AI not as an abstract technological artefact, but as a phenomenon embedded in history, culture, pedagogy, ethics, and social practice.

This volume insists that writing is not merely the production of text. Writing is thinking made visible. It is struggle, hesitation, revision, doubt, and intellectual risk crystallized into language. If AI replaces writing entirely, what we lose is not prose but the cognitive discipline that writing demands. Originality, in this sense, is indeed under threat—but not because AI writes. It is under threat if humans stop thinking.

AI exposes an uncomfortable truth: much of what passes as writing was already formulaic, derivative, and routinized even before machines learned to do it better. The real casualty of our era may not be originality itself, but complacency masquerading as creativity. AI forces scholars to ask harder questions of themselves: What do I think that a machine cannot? What ethical judgment, historical sensitivity, or lived insight do I bring to this text?

This edited volume does not mourn the past nor worship the future. Instead, it offers a space for critical reckoning. It invites readers to engage AI not as an enemy to be resisted nor as a saviour to be embraced uncritically, but as a profound intellectual challenge that demands humility, rigour, and imagination. If the age of AI marks the end of unthinking writing, it may also mark the beginning of more honest, reflective, and self-aware scholarship.

In that sense, this volume is not merely about Artificial Intelligence. It is about the future of thinking itself.

Saroj Kumar Rath
Abhishek Priyadarshi
New Delhi

Acknowledgements

To acknowledge this Symposium and its resulting edited volume is to admit, candidly, a curious mixture of trials, triumphs and missteps. We embarked upon an uncertain yet richly eventful journey—one marked by academic aspiration, institutional paranoia, and occasional mismatches of skill within the Indian academic ecosystem. That a small group of faculty members from the History Department of the College of Vocational Studies, University of Delhi, should conceive and execute a Symposium on Artificial Intelligence, and subsequently produce an edited volume with an international publisher, was both improbable and wholly unexpected.

When the Call for Papers was circulated to invite scholarly participation, the first objection reached us through Dr. Priti Roy, Assistant Professor of Commerce. She conveyed the response of a prospective participant—a faculty member from the Indian Institute of Management, Indore, Madhya Pradesh. The communication was blunt, even dismissive: he remarked that ‘we are animals and do not know how to think,’ and went on to suggest that, being trained in History, we were ill-suited to organize such a Symposium on Artificial Intelligence. We responded with measured irony, acknowledging—tongue firmly in cheek—that while we might indeed qualify as ‘animals,’ our scholarly menagerie still lacked one species—a Donkey—which he was welcome to supply should he choose to join us. Having thus discharged the moment with humour, we nevertheless received his observations with due humility and continued undeterred in our academic pursuit

But how did we decide on this topic?

On a foggy wintry morning in January 2025, we (Abhishek Priyadarshi and Saroj Kumar Rath) were discussing organizing a real academic event

on ‘Zero’, its authorship, its history, particularly through the prism of the mysterious volume of Leelawati – the medievalic magnum opus on mathematics. Immediately, the ever gullible Abhishek Priyadarshi picked up his mobile and called the erudite mathematician and legendary former Vice-Chancellor of the University of Delhi, Prof. Dinesh Singh. Without losing time, Saroj Kumar Rath picked up his mobile and started video-recording the conversation. Professor Dinesh Singh, in his usual flamboyance and disdain for amateurish activities, shamed us sufficiently to be chauvinistic before advising on how to organize a symposium and on its topic. It was during this discussion that Prof. Singh asked us to work on Artificial Intelligence and organize the symposium on this theme.

A few days later, we gate-crashed Prof. Singh’s abode at Greater Kailash, New Delhi. During the five-hour volley, when he fed us with food and drink, we argued, debated, discussed, deliberated and tried to convince him of our ideas based on his suggestion. The doyen of Indian mathematics refused to budge. Finally, we picked up the mobile phone and played before him the recording where he personally consented to his commitment to chaperone the symposium. Prof. Singh had no other option but to guide us in our strive to organize a far-reaching event. To further cajole him and to reduce the ferocity of direct fire from Prof. Singh during the course of the event, we roped in Mrs. Nilanjana Singh, the shadow life-partner of Prof. Singh. It worked well for us. Prof. Singh remained the nucleus and powerful guiding force with his unusual academic aura during the course of the Symposium. He has been, and still is, the guardian of our collective conscience.

But the symposium was not without *Delhi-Belly, the proverbial Bollywood satire on life in Delhi*, type difficulties. To smoothly organize the event, we created a pool of 20 faculty members from the College of Vocational Studies, University of Delhi. Except for three faculties (Vasundhara Chaudhary, Navin Ray and Prateek Kashyap), the rest of them never turned up.

We recalibrated our strategy and resolved to constitute a dedicated volunteer corps drawn from both current and former students of the College of Vocational Studies. As an initial gesture of camaraderie and trust-building, we hosted the group at several well-known restaurants along Hudson Lane in Delhi University’s North Campus, ensuring that all

fourteen volunteers—Rachita Singh, Parakh Shrinag, Srishti Sharma, Garv Gera, Chirag Sachdeva, Ayush Pandey, Lakshya Mehta, Anurag Gupta, Riddhi Bhardwaj, Sahaj Thakur, Suyash Shrivastava, Priyanjana Ghosh, Atharv K. Arya, and Aditya Nair—were generously hosted. For a fleeting moment, we wondered—half in jest—whether *ham saanpon ko doodh pila to nhi rahe hain?*, invoking the colloquial idiom of misplaced generosity literally translated into *feeding those very snakes, who ultimately bite the feeders*. Time, however, decisively proved otherwise.

As the Symposium unfolded, this group transformed into a formidable collective—an extraordinary convergence of courage, stamina, creativity, and unwavering commitment. They emerged as the public face and first line of engagement of the Symposium, consistently exceeding the responsibilities entrusted to them. Their firm yet courteous management of the proceedings inspired both awe and admiration, leaving an indelible impression on all who witnessed their professionalism.

For the College of Vocational Studies—a constituent college of the University of Delhi directly governed by the University—this high-impact academic event was a first in its institutional history. Its successful realization owed much to the academic leadership of the newly incumbent Principal, Prof. Shiv Kumar Sahdev, himself a mathematician. Prof. Sahdev’s leadership was marked by an exceptional attention to detail and an uncompromising commitment to academic excellence. During the inauguration of the Symposium’s official logo, he graciously agreed to preside over a modestly organized ceremony. While inaugurating the logo, however, he candidly rejected it on the spot, citing its technical inadequacies. The logo inauguration ceremony, which we had hoped would be a high-voltage success, ultimately turned out to be an anti-climax. The logo was subsequently redesigned and replaced with a more appropriate and visually compelling version. Throughout the Symposium, Prof. Sahdev remained the true patron and guiding presence—a familial head in spirit and substance—embodying a model of academic leadership worthy of emulation by contemporary academic administrators.

The travails continue, as merely a few days before the Symposium, we got a mail from Prof. Snezana Lawrence from London, informing us that her visa was rejected by the Indian Embassy in London because we (the organizers) failed to provide the clearance letter from the Ministry of

External Affairs. We kept this news to ourselves and quietly acquired the clearance from the MEA and persuaded Snazena to apply afresh. This time, her visa was approved and she delivered a flamboyant speech at the Symposium.

The Symposium received academic support from three continents: Asia, of course, Europe, and North America. Scholars from all over India gravitated to the Symposium to deliberate on Artificial Intelligence's efficacy in social science and beyond. The Symposium featured three distinct formats of participation: keynote addresses, invited extempore presentations, and rigorously researched studies selected through a call for papers on specific themes. From these three formats, twenty-three papers were curated for inclusion in the conference proceedings.

We gracefully acknowledge the generous support of the Indian Council for Social Science Research (ICSSR), Ministry of Education, Canara Bank, and VidyaAI for sponsoring the event. Prof. Dhananjay Singh, Member Secretary of ICSSR, remained a stellar support throughout the Symposium. We also wish to acknowledge the invaluable behind-the-scenes contributions of numerous anonymous reviewers, dedicated supporters, and colleagues from the organizing committee (Ankit Kumar, Saurabh Kumar Shanu, Vikram Raj, and Sumit Kumar).

In a moment emblematic of scholarly devotion, Saurabh Kumar Shanu appeared for the defence of his Ph.D. degree during the course of the Symposium itself. In a scene both unconventional and unforgettable, he defended his thesis beneath a secluded staircase at the India Habitat Centre—the very venue of the Symposium—transforming an overlooked corner into a sanctum of academic resolve. Standing sentinel was Ankit Kumar, who assumed the role of watchful guard, ensuring that no onlookers breached the sanctity of the moment or disrupted Shanu's concentration before the unblinking gaze of his Ph.D. degree assessors in his laptop camera. Perhaps the most remarkable aspect of the Symposium was that Shanu emerged with his doctoral degree forged amid the bustle, smoke, and scholarly ferment of the event itself—a testament to the extraordinary commitment, adaptability, and quiet heroism of the organizing committee.

Their collective efforts ensured that the event remained rigorously academic, intellectually introspective, and highly productive, yielding tangible outcomes. We trust that this volume will resonate with discerning

general readers, curious observers of artificial intelligence, and the wider intelligentsia.

Saroj Kumar Rath
Abhishek Priyadarshi
14 January 2026
Makar Sakranti

About the Editors

Saroj Kumar Rath, PhD, has written extensively on the history of India's agonizing over rival powers. His history book, *Fragile Frontiers: The Secret History of Mumbai Terror Attacks* (Routledge, 2014), has been used as course material by leading universities in Europe, Asia, Australia and North America. *Fragile Frontiers* is adjudged as *Best Terrorism Books of All Time*. Saroj has edited two volumes on cyber security in India, *India's Cyber Security* and *Securing Digital India*, published in 2015 and 2016. He is a visiting Professor at the Universidad Autónoma de Nuevo León in Monterrey, Mexico, and an Assistant Professor of History at the University of Delhi. From 2008 to 2013, he worked as an Assistant Professor at Hosei University, Tokyo. Before that, he had a brief stint at Deakin University in Melbourne, Australia, from 2007 to 2008.

Abhishek Priyadarshi, PhD, is currently serving as an Assistant Professor at the College of Vocational Studies, University of Delhi, a position he has held since 2023. Before this, he held teaching positions across several prestigious colleges within the University of Delhi. Before transitioning into academia, Abhishek was an ace Assistant Commandant in the Central Reserve Police Force under the Ministry of Home Affairs, Government of India, from 2012 to 2014 as a leader of paramilitary where he devoted his time to the strategic security of India. His forthcoming book, titled *Partition and the Congress Muslim League*, is set to contribute significantly to the discourse on Partition and communal politics in pre-independent India. He has successfully received grants from the Indian Council for Social Science Research, Ministry of Education, and Indian Council for Historical Research to organize national and international symposiums.

Organizing Committee

Prof. (Dr.) Shiv Kumar Sahdev

Principal, College of Vocational Studies

Prof. Arun Kumar Acharya

Universidad Autónoma de Nuevo León, Monterrey, Mexico & Department of Anthropology, Sambalpur University, India

Dr. Manoj Kumar Gupta

Senior Research Scientist at Vertex Pharmaceuticals, USA, formerly at Joslin Diabetes Centre, Harvard Medical School, Harvard University, Boston, USA

Prof. Manoj Kumar Dora

Director, Centre for Intelligent Supply Chains, Anglia Ruskin University, Cambridge, United Kingdom

Prof. Mahmoud Masaeli

International Development and Global Studies, University of Ottawa, Ottawa, Canada

First Keynote Address on Artificial Intelligence, Re-examining the Social Sciences through the Lens of AI & IT Innovative Intersections: Charting the Future of Social Inquiry in the Digital Age

Dinesh Singh

Former Vice Chancellor, University of Delhi

DOI: [10.1201/9781042018116-1](https://doi.org/10.1201/9781042018116-1)

We began the ceremony by reciting from the Rig Veda. And I cannot help but recall one of the openings of the Rig Veda, *Anu bhadrah kratavo yantu viśvataḥ*. Let noble thoughts come to us from every direction, and those who represent different parts of the world, who are present here, symbolize that thought from the Rigveda. So, I welcome that.

Let me confess at the beginning that I'm not really an expert in artificial intelligence, but I've been dabbling with it for about two years now, and I want to share my experiences. I'm not sure how many of you are aware of the seeds of artificial intelligence, since this is a conference on the social sciences, but I must also say I have a bit of a disagreement with this business of the division of disciplines. You know, we never had this in India, and we set up universities in India much before the West even knew what a university was. I speak of Taxila, and we didn't have such disciplines that so and so is from history, so and so is from biology, so and so is from mathematics. There was nothing like that. We were completely trans-disciplinary in nature. We were immersed. And I'll give you one example, around 200 years before the time of Jesus Christ, in India, profound

mathematics came out of Sanskrit poetry, profound mathematics. Mathematics was reinvented in Europe 1800 years later.

Why did that happen? Because we were good with data. We were good with combinatorics. We were good with mathematics, and we were good with the social sciences, all embedded in an individual, and that we have forgotten. The Whites came and taught us that this is history, this is mathematics, separate them, and we divided them. And now it's too late. But I still address the young here to think in a transdisciplinary way. The national education policy shouts from the rooftop repeatedly that think in a transdisciplinary way, not interdisciplinary. I come from Banaras, and there is a saying there that one never knows which seed will bear fruit from which soil. This wisdom applies equally to education, to research, and to transdisciplinary pursuits.

You never know what discipline will give rise to fruits; in what discipline let India learn that a little bit. I'm happy that we have chosen artificial intelligence and the social sciences that will help the young here to see things in a slightly different perspective. The roots of this business of artificial intelligence go back a long time, and I want the young to understand this. I am primarily addressing the younger generation here. You know, I'm passing on wisdom to you. What I am trying to say is this Jacquard loom comes to my mind, it is one of the first instances of the roots and seeds of artificial intelligence. In India, the tradition of weavers is very old; the Indian tradition of weavers is ancient and dates back to before the time of Christ. But I am so disappointed that India could not invent a mechanical system like the Jacquard loom. We could not; I want the young to think about it. The idea of an algorithm comes from India.

The word algorithm was invented to describe processes that were unique to Indian mathematics, and yet we could not invent the Jacquard loom. Think about it. We could not invent the telescope. We could not invent the microscope. Why? I want you to ask this question to yourselves. We invented marvellous things in mathematics. We invented calculus 500 years before Newton. Bhaskaracharya in the 12th century, invented calculus to study planetary motion. All those theoretical advances were there, but for some reason, we could not apply that to develop technology. I want you to think about this. So, the Jacquard loom is the first instance that comes to my mind in history where there is something called automated processes. Then

there is Charles Babbage, who was the first man in the history of mankind to create a computer. He was a genius, 100 years ahead of his time. Again, no connection with India. India was totally absent.

India has made significant progress and achieved remarkable accomplishments in many fields. We developed mathematics through Sanskrit texts nearly two centuries before Christ, which is unparalleled. However, in areas where we have not been able to advance as far, it is essential for the younger generation to recognize the need for greater focus and effort, as neglecting these fields could have serious consequences for the country.

Just ask yourself this question. Then we move ahead. Artificial Intelligence starts developing with the advent of computers. I have a friend, a very distinguished mathematician. Unfortunately, he has now almost lost his mind. He's in America. In 1986 when I was just starting out as a mathematician, he gave me a demonstration on his personal computer. He just purchased a Mac. And the Mac of those days was very different from the Macs that we see now, but for that time, it was a great machine, and he showed me how he had invented, for that time, an artificial intelligence platform called Graffiti. The New York Times carried the article on that also, what did Graffiti do? Graffiti was inventing conjectures in mathematics, and what was the discipline of mathematics, graph theory. This is 1986, since 86 till about late 90s, I kept shouting from the rooftops in India. Why is India not doing something related to Graffiti? There are so many graph theories in India, and I respect all of them but not one of them looked at Graffiti. I have it from Paul Erdos, one of the greatest mathematicians of the last century, he personally told me that he was busy trying to solve conjectures that graffiti was generating. That's how important graffiti was. But India slept blissfully in this field.

Graffiti generated deep and hard conjectures. How did he do that? Because the software in the Mac allowed my friend, Siemion Fajtlowicz was his name, this mathematician to write codes in a complicated way, to look at huge amounts of data, to see patterns, and then to guess. And that's a bit of artificial intelligence, one of the first instances I've also written about this. But, you know, I don't know how many people read what I write, so I decided I have a captive audience here today, I'll tell you about that.

Artificial Intelligence has existed for a long time in our lives. When the newspapers and media announced the GPT 4 or ChatGPT, that's how it started. But you know, you are already using artificial intelligence. Most of us, at least in urban areas, when we decide to make a journey by road, we always use Google Maps. Well, that uses artificial intelligence. It uses data, real time data, and a bit of artificial intelligence, and that's how it tells you that you will get from one place to the other in so much time, and you will encounter so much traffic on that route. And by the way, that is also predicting the future for you most of the time. Just think about it. There is no astrology here, but with simple mathematics and computational science, you can even predict the future.

I can predict today, how many years from now there will be a lunar eclipse. I can tell you exactly where, day after or a year from now, Saturn will be, where the moon will be. I can make that prediction today with 100% accuracy, and I don't need any astrology. That is the power of ideas. And I want the Indian youth to come to grips with these ideas now that you have artificial intelligence at your disposal, you can do that in a very effective way. And I'll try and demonstrate very quickly to you. Also, one has to be careful.

As Dhananjay also said, you know that Shri Krishna told Arjun, verify your sources, verify where you're getting your data from. You have to be very careful. And let me narrate a personal experience. You know, my area of research involves a term called operator mathematical research, and I communicate by email with my collaborators who are in the United States. And this collaboration happens very frequently, and I use email. And when I started this, and after a certain point some agency sent emails and called me stating that we have a brilliant job for you, 15,000 rupees a month as a desktop operator. And I hadn't applied for anything as such. So, you understand sometimes artificial intelligence can go absolutely haywire.

It just picked up the term operator and then focused on it completely. And so you have to be alert all the time. And there are many things that you could do with it, but there are many things that you should see how to avoid. You must understand that this is about social sciences. The social scientists tend to stay away from technology, from science. You could say the same about scientists and science. They stay away. That is why I bemoan this fact that we have made these divisions. India did not have

these divisions in the past. We have blindly adopted what the West has given us, whether it was suited to us or not, but the national education policy, as I said earlier, shouts from the rooftops to avoid this partitioning, to make sure that you're not in silos, to make sure that what you learn is transdisciplinary.

So think about that, but artificial intelligence can help make it happen in so many ways. Let me illustrate. Some time back I asked a well-known AI platform to generate four lines of poetry in Hindi similar to what Maithlisharan Gupta used to write.

Maithlisharan Gupta was a national poet, his poems were of very high standard and you should read Jayadrath Vadh, and his other works. So, I asked the AI to write 4 lines in Gupta's writing style. And I asked another AI platform who had written this poem and it quickly attributed that poem to Maithlisharan Gupta.

Do you see the power?

Do you see what you could do? Do you see the possibilities? I don't know how many people here know that in the English language, the letter E occurs with the greatest frequency. It is a fact of the English language. Take any passage of the English language, count the frequency of the letter E and all other letters and E will be way above all the other letters. I don't know how many people in the social sciences know this or ever use this to teach something to students. You can do so much, but the social sciences tend to stay away from data. Now, with artificial intelligence, handling data will be very easy for you. When the whole platform was announced. I downloaded many of them and I asked them any questions related to simple data. I gave them an array of numbers. I said, arrange them in increasing order in half a second. It did that. I said, find the mean median mode. It answered that. Then I asked which numbers are multiples of which number? This is a large list of numbers and it told me that a certain number is divisible by that, and this is not. Do you realize what you have in your hand? You don't need to be good at mathematics. You just need to harness that kind of power with IT, because this conference is about IT and artificial intelligence, and they go hand in hand with it. You could do a million other things if you use AI with IT, which you have to use, otherwise, it doesn't work. Some years ago, the Deputy Chief Election Commissioner of India, an IAS officer, played around with electoral roles. He had all of India's electoral roads. This is

long ago, so please don't look for any controversy here; it's a very non-controversial act. And you know what he did? He looked at all people whose surname was Arora, and because he was versed with software, a little bit of AI, he wrote a quote to see which city in India has how many Aroras, population wise? And you will not believe it. He could see the migration pattern of Aroras from Pakistan after partition. So there is a border town where there is the biggest concentration of the Arora community. Then they move in three directions. And as they move away from this border town, the population diminishes. Then you can see, this is the migration pattern, and I don't see anyone in my university, from humanities or computer sciences, even read Balakrishnan's research paper or even know about him.

I'm telling you, I don't know of any historian, I don't know of any sociologist, I don't know of any political scientist who has even bothered to read the book that Balakrishnan published. It is a tragedy we have so much in our silos. He also did another exercise. He looked at common place names in Sindh and Tamil Nadu. And you're not going to believe this. It reveals so many mysteries. He found 12 place names in Sindh, Pakistan, which are exactly identical to 12 place names in Tamil Nadu. Why did this happen?

Think now how many places in America have been named after places in Europe. They use the same names in America. There is a London in America. So you have all these things. And either there was a migration from Sindh into Tamil Nadu, or from Tamil Nadu into Sindh. It is a problem for historians and others. No one has bothered to follow this stuff.

Incidentally, there is a part of Balochistan where they speak a language called Ruhi, which is very similar to Tamil. I see very little work being done on that, but I want you to understand that artificial intelligence and AI will affect you in a million ways, and you could do that in your own but play around with it. It's not going to bite. You spend a little money, get the advanced versions, and then you can do magic with them. Look at Salman Khan of the Khan Academy. He does amazing things with artificial intelligence. He has collaborated with GPT four, and he has got this platform called Khan amigo and students learn in a personalized way.

When we were students, I wished some good teacher would teach us one on one, so that we could pace ourselves and then learn according to our own

habits, desires and abilities. GPT four in Khan amigo allows you to do that, and I'm not just talking there is hard data to prove this. Khan has data to show that the pupils in his class who learn through Khan amigo perform far better than pupils who don't do that, their performances have shot up.

India needs to awake to this. India must awaken and humanities has to say this has nothing to do with technology. All your disciplines will be taught like this, and you could get your students to light up their minds and do other magical things. Khan amigo also does amazing things.

Khan shared the experience of his daughter, who was in grade eight or nine. She was writing an article on Scott Fitzgerald, the author of *The Great Gatsby*. And she had to write some comments on *The Great Gatsby*, and she got into a personal conversation with Khan amigos, version of GPT four. It was like she's talking to a human back and forth, back and forth. And she said this was a surreal experience. Due to her conversation with AI, she thought that Scott Fitzgerald, himself was talking to her, and she gained many insights.

I would like to point this out and request even Mr. Doja that when we employ teachers in our institutions, we only look at the academic qualifications of the person and don't look at their relation with AI. But it is important that we add the perspective of artificial intelligence, otherwise, India will be in deep trouble. There are three things that will dominate the development of ideas, technology and our society in this century: nuclear fusion, quantum computing, artificial intelligence with big data. In nuclear fusion, which is round the corner in less than five years from now, Microsoft or Google in America will start to get electricity from nuclear fusion. This is 100% clean energy and limitless energy.

Where does India stand in all of this? I looked into it and they are saying that we will conduct an experiment in 2060. I don't blame any government for this. I blame all of us. All of us who are connected in the university and knowledge systems are so blissfully unaware of anything that exists around us. We are so blissfully unaware, happy in our own existence, not realizing that we are betraying the generations that come to us with faith. We must awaken their minds. And how do you do that? Recently, Elon Musk released a video that has gone viral on the net. He has a humanoid using artificial intelligence in real time that does kung fu. It's as fluid, as fast, as effective as any human.

Where are we headed? I want India to awaken. When the Apple phone was designed there was nothing like science or technology, or humanities. All of them got together, people from the humanities, from design, from science, from technology, from mathematics, and they built the iPhone and the technology in it, and other things. I want you to understand this. I want you to take this into account and use AI to make your journey fluid, easy and productive. I personally have benefited in amazing ways, in my own mathematical research. Just a few days ago, I've been struggling with a mathematical problem with my collaborators for the last one year. I had solved all cases in that problem except one, and I was stuck there. And I asked my collaborators, and I asked some other experts across the world, nobody had an idea.

Then I asked Grok, Elon Musk's version of AI, the advanced version, and Grok immediately cited a mathematical theorem that was 110 years old, discovered by a distinguished Cambridge mathematician of that time. Not only that, it was published in a journal in 1914 in the Proceedings of the London Math Society, and it extracted that theorem for me. And it said, using this theorem it will be useful. Sure enough, I solved my problem, and I have a marvellous paper being published in a very well-known journal. Just because I used artificial intelligence, I want you to know it's your friend. Start making use of it. Start getting friendly with it, and you will see amazing things that will come out of use of these things. I have played around so much, and I've done so many things with it.

In languages, you can do marvellous things. You give passages and play around and say, compare this to that, which is better? Why is this more creative? But it will give you good answers. You don't have to accept that. It stimulates your mind. You don't have to learn the technology behind it.

That's another thing in India, we don't understand the neural networks, maths or programming behind it, our knowledge about it is almost negligent. China is way ahead of us in this. China is as powerful as America in Artificial Intelligence. I've seen their humanoids play basketball or football, I think one of the two games.

But you are young; the future belongs to you. Wake up to this call. Take advantage of this conference. This is just a beginning. The idea is to spark your minds and to then show the right path with the torches that the sparks in your minds will light for the well-being of this nation.

Second Keynote Address on Artificial Intelligence, Re-examining the Social Sciences through the Lens of AI & IT Innovative Intersections: Charting the Future of Social Inquiry in the Digital Age

Dhananjay Singh

Member Secretary, ICSSR, Ministry of Education, Government of India

DOI: [10.1201/9781042018116-2](https://doi.org/10.1201/9781042018116-2)

It is a great privilege and pleasure for me to speak here today, especially in the presence of Professor Dinesh Singh, whom I have looked up to since my student days at JNU, and Guruji. Even without this honour, this would have been a momentous occasion for me because of my long-standing interest in artificial intelligence. The experience of being here with Professor Dinesh Singh and Guruji is something that perhaps no simulation of artificial intelligence can articulate, for faith and devotion (shraddha) are integral to the human core. The human element must remain at the centre of technology as well.

As many of you know, the term ‘artificial intelligence’ was coined by a young assistant professor, John McCarthy, on June 18, 1956. He introduced this term in a proposal for a workshop much like this one. In that Rockefeller Foundation proposal, he wrote:

‘Every aspect of learning or any other feature of intelligence can, in principle, be so precisely described that a machine can be made to simulate it. An attempt will be made to find out how to make

machines use language, form abstractions and concepts, solve the kinds of problems now reserved for humans, and improve themselves.'

This historic vision of McCarthy's guides my thinking today. Here, what I intend to talk about is AI and the social sciences: how the big data boom and AI simulations of social and human behaviour are transforming social science disciplines and research production, and their applications. I must admit, an AI model or robot could perhaps articulate these ideas even better than I can right now. But why I am here, speaking these ten minutes, is because I lead the Social Science Research Council of India. We have funded and shaped many research grants and concepts in this area—on AI and women's empowerment, AI's impact on employment, AI and ethics, and so on.

Yesterday, for the first time, I tried using ChatGPT. Why I did so is an interesting story of its own in my personal experience at ICSSR. But suffice it to say, I asked ChatGPT: *"If someone asks why they should listen to you (an AI model) instead of listening to me, what would you say?"* ChatGPT answered in effect: "You should consider listening to me if you want a well-informed, fast, and neutral synthesis of ideas drawn from a wide range of academic sources, theories, and data. I have been trained on a large body of peer-reviewed literature, textbooks, and expert commentary across sociology, psychology, political science, economics, and more, up to [the year 2024]. You teach English and know a bit of social science and Indian philosophy, but I am neutral, whereas you might have your own worldview or ideology." In short, ChatGPT was saying, "I'm at no risk of personal bias; I can provide clarity and context."

That response made me feel, humorously, that I'm at great risk of losing my job. So I asked myself: *What does the Bhagavad Gita say about knowledge?* The Gita reminds us that "knowledge is a great purifier." In our Indian tradition, vijñān (scientific knowledge or discernment) is as important as jñān (wisdom or philosophical knowledge). I told myself: let me start by saying that. Knowledge isn't just about accumulating ideas and concepts in external storage (like textbooks and data banks); knowledge (gyan) is a deep, reflective internalization of what we study, and it has an

important element of ethics. Perhaps if I emphasize this, I still have a chance to make my point.

Then I asked ChatGPT again about its own limitations. ChatGPT replied: *“I’m not a human expert. I do not generate new research. I do not interpret data with personal bias or lived experience. What I offer is breadth, clarity, and context.”* That response pointed me to an important insight: it suggests the necessity of collaboration. Social scientists and natural scientists must work together; collaboration is the foundation of research. Research cannot happen in silos. We have been trying to convince our colleagues in the natural sciences that social science is integral to studies on climate, energy, agriculture, and technology, because understanding lived human experience is crucial in all these fields.

That brings me to a fundamental question: How long will it take for AI or technology to truly capture the human aspect? When John McCarthy applied to the Rockefeller Foundation, he asked for \$13,500. He was given \$7,500—almost half—because at that time the foundation was not convinced of the great service AI could do for humanity. Today, the human element remains an open question, yet new possibilities arise: personalized data, interpretation, and the generation of new insights. But how long can those possibilities sustain? That is a big question for all of us.

ChatGPT further answered my questions by saying: *“Well, since I’m not a human expert, what I can do is help explain theories, trace the history of ideas, compare perspectives, flag biases, and even critique methodologies.”* In other words, it's up to us, as social scientists, how we prefer to use AI at different stages of our research. The entire ecosystem of research is going to change, and the very concept of originality is going to change with it.

Let me give you an illustration. Imagine I am Arjuna on the battlefield of the Mahabharata and AI (ChatGPT) is Krishna. AI tells me, “Decide for yourself what weight to give my words; always cross-check them with primary sources, recent peer-reviewed studies, and the scholars active in the field.” In other words, we should not take AI's word unquestioningly. Therein lies our scope and our role. We are not in opposition to AI; we just have anxieties about it. Actually, we must consider how the human world and the AI world will coexist. Some scientists even say that in the future, the human sphere of influence in certain domains may shrink. But what will keep humans relevant? At least so far, in social science research, we have

been using AI as a tool—using large language models to assist us. India, in particular, has great scope in this endeavor.

For example, I have the additional responsibility of heading the Commission for Scientific and Technical Terminology, where we develop scientific and technical terms in 22 Indian languages across about 17–18 disciplines (including tribal languages). This itself shows the big potential for India to develop our own models of machine learning and ICT for education. In education, AI is actually bringing back the guru–shishya tradition in a new way: learning is becoming more personalized. Assessments and evaluations will become more automated; problem-solving will become more creative; and virtual simulation will make the learning experience in education much richer and more effective than it was.

Likewise, in research, I believe we will see much more democratization of the process. You won't have to depend solely on a single “guru” or supervisor as in the past. I want to share an anecdote: Two days ago at the India Habitat Centre, I met a colleague who used to be my hostel roommate at JNU and now heads a premier institute. He told me how, when he arrived at JNU in 2001, he could barely speak English and didn't even have a computer. An old professor (fresh back from Canada) said to him, “I want to meet you in Hindi,” but then curtly advised him to “write me an email and ask for an appointment.” My friend was thinking, “Sir, it's 2001, I don't have a computer, I only barely speak English—am I supposed to dash to the computer lab, compose an email, and hope to get a reply in a day or two?” The point is, in those days these barriers existed. With AI and modern tech, such problems become easier to solve.

My daughter is in 9th grade, and I recently wrote an article on the fifth anniversary of India's New Education Policy. I talked about the rise of what I called “new natural intelligence” (the integration of AI) alongside the NEP. We educators and teachers must wake up to this new reality: if we do not use technology and do not act as true facilitators of learning, students have countless other options—often better ones. These are challenging times, but with that challenge comes greater responsibility. We must learn from AI and develop the competencies it demands of us.

As I said, social science research is extremely crucial for national development. If India (Bharat) is to be fully developed by 2047, social

science must play a big role. Economic and socio-economic transformation cannot happen without understanding society. That is why even health experts now talk about “One Health”—the idea that the health of humans, animals, and our environment are interconnected. It is the perspective from the humanities and social sciences that underlies this holistic concept.

So, in summary: AI offers us enormous possibilities, and we should harness these possibilities to develop new ideas. The very notion of originality in research is going to change. We are in an era where old and new ways of knowing converge. I wanted to refer to some classical concepts like śabdāhāraṇa and ārdhāhāraṇa (from Sanskrit tradition on learning), but to put it simply: the notion of gaining knowledge through words and listening is not fundamentally different from how AI processes language and data today. As researchers, we have to adapt to these new paradigms.

AI can certainly give us ideas and even predict new social phenomena to study. For reviewing existing research in a field, AI can do it far faster than we can. It could shave off a year or more that our MPhil or PhD students used to spend reading everything from scratch. But can AI replace reading all primary and secondary sources ourselves? It can summarize them, yes, but new research emerges from finding gaps and inconsistencies in the literature, and that requires deep, personal engagement. Only by reading thoroughly and comparing studies can we spot those gaps.

Of course, even beyond AI, digitization and ICT have already revolutionized research. Hypertext, search engines, and online databases have changed how we work; they make research more accessible than ever.

As AI advances, I find it interesting that some draw analogies from our ancient philosophy. For example, one commentator compares Patanjali's Yoga-sūtras on the mind to modern data concepts: equating buddhi (intellect) to our big-data infrastructure, aham (ego/self) to contextualized data, and manas (individual mind) to personalized data. This metaphor suggests that if AI were to develop its own consciousness, it might foresee our responses. If that happens—if AI can pretty well predict all our actions—then where will we human beings fit in? We could no longer simply “go back to our offices” as before if AI can do our work. Then what will we do?

Consider self-driving cars: Stanford (where McCarthy worked) saw the first self-driving car. Just recently, there was an incident where a self-

driving car made a huge wrong U-turn and caused an accident. These issues will improve over time, as they always do. I see two possibilities moving forward—one material and one spiritual. We must not panic. Remember when Elon Musk launched a Tesla Roadster into space? The world didn't panic then, and similarly we must not panic at the prospect of AI. Technology fundamentally is built for humans – humans are the end-users. Otherwise, we'd end up like in *Star Wars*. Since humans will remain the end-users, and since market forces (driven by powerful corporations) will shape these technologies, we have to ensure that ethics and compassion guide us. These human values are what will make our future sustainable as we contemplate the future of humanity.

Finally, on a personal note: I have always been interested in Indian philosophy (and even Greek philosophy), and I believe philosophy will be extremely important in this era of AI—it already is. For instance, UNESCO held a conference on AI and ethics in Slovenia recently, and I had the privilege of participating. I think more and more of us—scientists, researchers, scholars—should listen to philosophers as we build a community to coexist with AI in our work.

Third Keynote Address on From Papyrus to AI: A Historical Perspective on Tools in Mathematics

Snezana Lawrence

*Department of Mathematics and Design Engineering, Middlesex University,
UK*

DOI: [10.1201/9781042018116-3](https://doi.org/10.1201/9781042018116-3)

I'm going to talk about the historical perspective of how tools have changed the way people think and do mathematics. I think this is important because mathematics is a kind of universal language that even applies to fields like agriculture. Apart from being a mathematics historian, I think it still has a lot of lessons for people in other disciplines as well. So what this talk is about is given here (on the slide), and I want to concentrate on several points: the broader relationship between mathematics and AI. We focus on that relationship because it tells us something about new tools and how the practice in various fields influences the development of those fields. Looking at history, we can draw some conclusions at the end about the future, for example, what I think is important when introducing AI in all aspects of education, research, and life.

I've chosen three episodes from the history of mathematics to talk about. This is really intellectual history as well. As I said, I see mathematics as primarily an abstract way of thinking that applies to every culture and context. The episodes I picked to illustrate this are: the invention and use of papyrus, the invention and use of the mechanical printing press, and the invention (or use) of computers in doing mathematics. These are basically the three big phases I want to cover. The relationship between mathematics

and AI can be thought of in three different aspects. First, mathematics as a computing discipline – something we use in everyday life, from ancient times up to now. We use it for making things, measuring things, counting, and basically as a tool in all science and engineering. Second, there is research mathematics, which is about discovering new things, new ways of thinking, and defining those ways with new tools – essentially developing new tools for thinking. Third, there is the body of mathematical knowledge that exists, which is constantly being developed and needs testing, communicating, and disseminating. All three of these aspects are important when we talk about AI, and they are applicable to other disciplines too. Let's see how the three historical phases I mentioned unfolded.

Before papyrus – actually, before written history even began – there was still mathematics. We have things like the Ishango bone, which is about 20,000 years old, and it shows that there was mathematical thinking even back then. We know this because it had incisions that showed people already had ideas of, for example, prime numbers, and they knew how to add and subtract. They even had a primitive form of notation for those ideas. Later we have the Babylonian culture (around 2000 BC). They used clay tablets and already had quite developed mathematics. For example, they knew the relationships that were later (thousands of years later) called the Pythagorean relationships. But the clay tablets were heavy and hard to transport or store. So before papyrus, it was really difficult to preserve mathematical knowledge and share it with other cultures. I think that's the most important aspect: technology matters for how we share ideas.

Then came papyrus. With papyrus, mathematics could be written on a relatively small papyrus roll – light and portable, easy to take around and share. This is what happened at the Great Library of Alexandria (around 300 BC), where scholars collected as much mathematical knowledge as they could from the Greek world, copied it, and stored it on papyrus scrolls. From that point on, we could start to share mathematics between different cultures. In fact, symbols were already being developed to shortcut some of the longer calculations or proofs. Papyrus was used for many centuries until paper came along later.

One of the classic examples is Euclid's *Elements*. Euclid's *Elements* is a collection of 13 books written around 300 BC in Alexandria. It's the oldest large-scale mathematical work we still have. It compiled most of the

geometry known at the time. It starts from definitions (like a point, a line, etc.) and develops geometry all the way up to things like the Platonic solids (the three-dimensional shapes like cubes and tetrahedra). I chose *Elements* because it had a huge influence on mathematics worldwide and illustrates our points.

There is a very famous papyrus copy of *Elements* from the 3rd–4th century AD (so already several centuries after Euclid's time), and it shows that even 400 years later, people were still writing Euclid on papyrus. The picture on the slide shows a diagram from that copy. Notice how the diagram is drawn, and imagine that back then it was just explained in words on the papyrus. In modern textbooks we would describe the same relationships more explicitly with labels. It took another ten centuries or more for mathematics to be described and explained in the more formal way we use today, but the mathematics in Euclid's book remained essentially the same through all that time.

Eventually, mathematics moved to paper and print. The first printed edition of Euclid's *Elements* was produced by Erhard Ratdolt in 1482. It was based on a Latin translation of an Arabic version (the Arabic version had been translated by a scholar named Campanus of Novara). Ratdolt used creative techniques – bent metal lines and gold leaf – to print the geometric figures nicely. With printing technology moving so fast, Ratdolt's edition was distributed across Europe. This also made people start to wonder if there were any gaps or mistakes in the knowledge being passed around.

Another big development was that mathematics became available to the general public. By the 16th century, scholars started printing Euclid's *Elements* in vernacular languages, not just in Latin. The copy I show here is the 1570 English edition translated by Henry Billingsley, which has a famous preface by John Dee. I include it because it even has pop-up diagrams! (Mathematicians and book historians find those fascinating.) In any case, by that time, people around the world could read the same mathematical work in their own language.

So by the printing-press era, mathematics was being written, copied, and printed on paper, and knowledge was spreading widely. In the 19th century there were more developments (like people critically examining Euclid's axioms), but I won't go into them here for lack of time. The main point is

that by the 19th century, mathematics was becoming very deeply studied and shared.

Then we come to the 20th century, which brought the invention of computers. At first, computers seemed like just fast calculators, and people didn't think they would change mathematics much. But then some discoveries showed otherwise. I think the crucial moment was realizing that computers can do much more than crunch numbers – they can help *invent* new mathematics. One famous example is the invention of fractals. Benoît Mandelbrot, who was working at IBM in the 1970s and 80s, used computers to explore ideas about roughness and dimension. He took some patterns and theories from the 1920s (which were hand-wavy) and tested them with computers. Around 1980 he essentially discovered the mathematics of fractals (the Mandelbrot set, Julia sets, etc.). This was a new kind of geometry that came from using computers to explore patterns. It shows how computers enabled mathematicians to see something new that was hard to guess by hand.

Another big change in the computer age was TeX/LaTeX, a computer program (created by Donald Knuth) that everyone can use to write mathematics professionally on a computer. This might sound technical, but it's extremely important: LaTeX allowed mathematicians around the world to easily write and share papers. Even more important was the internet: mathematicians began sharing their papers on the web (think of sites like arXiv). Many top mathematicians started posting preprints online instead of only through journals. This means mathematical knowledge got distributed worldwide faster than ever. It would be interesting to study how quickly information now flows among mathematicians; perhaps AI tools could even analyze that exchange and spot new patterns of collaboration.

Now, bringing us to the present – AI and mathematics. Just a few years ago, many people thought AI wouldn't work for mathematics. Large language models (LLMs) like the early versions of GPT didn't seem to handle math well at all. But we've seen enormous progress very recently. Today there are specialized systems like DeepMind's *AlphaGeometry* and *AlphaProof*, designed for math problems, and even something called *Harmonic Aristotle* (a system trained on mathematical theories). Even general AI like ChatGPT can do surprisingly well on simple math problems

and can teach low-level math. For example, if you ask ChatGPT simple math questions, often it gives correct answers or helpful explanations.

What can we use AI for in mathematics? Looking back at the other historical changes, we saw that new tools didn't immediately change the essence of mathematics, but they changed *how* we do it: how we distribute and question theories. We can apply that idea with AI. One thing AI might do is help *find gaps* in our existing knowledge – essentially doing the work of searching for patterns and conjectures that might take humans centuries to spot. For example, a machine learning algorithm could comb through thousands of papers or data and suggest a new hypothesis. Also, AI can help make other tools easier to use. Suppose we have something like AlphaGeometry that is very powerful but also very specialized; if you're not a geometry expert, you might use ChatGPT to explain what AlphaGeometry is doing in plain language. In the same way, we can use proof assistants (like Coq, Lean, or Isabelle) in combination with AI: the assistant can check and verify proofs, and AI can help guide how to use those assistants.

I think one of the most hopeful aspects is sharing and collaboration. With AI, by definition, you can share your questions and ideas with a machine that has been trained on so much information. In a sense, you are sharing your knowledge with the world (via the model) and getting collective knowledge back. This sharing of information can bring new discoveries to light. It's like a new way of collaborating: you might say that your knowledge is part of what the AI “knows,” and when it gives answers or suggests ideas, it's reflecting that shared knowledge.

Of course, there are problems to consider. One big problem is that AI is highly commercialized. The companies developing AI have huge incentives to keep their technology proprietary. We're already seeing signs of this: things that once were freely available can become paid services. We saw something similar 20 or 30 years ago: at first a lot of data was on the web for free, and then companies started charging for access or locking it up with copyright. So it's important to be aware of this. If AI tools become closed off or expensive, they could become tools of disengagement – meaning people who can't pay or have limited access might be left behind. We have to influence these developments so AI doesn't become a source of inequality.

Another point is about creativity and hunches. Historically, big mathematical ideas often came from human intuition, hunches, and even from mistakes or dead-ends. For example, sometimes something that looks like a failed idea (a byway) later becomes important in a new context. It often takes decades or centuries for a mathematical insight to pay off. AI is very powerful, but right now it doesn't really make creative hunches on its own. So we shouldn't forget the human side: humans generate new ideas through intuition, imagination, and serendipity. My suggestion is: don't let AI take the lead completely. It can help, but human creativity is still key. For instance, if you have a conversation with an AI about a math problem, you might get interesting ideas back, but ultimately you should trust your mathematical intuition and check things carefully. Hunches and creative leaps are hard for AI to reproduce – that's something we still bring to the table.

So to wrap up, AI is a powerful mathematical tool, but we should think of it as one more tool in a long line of tools (like papyrus, printing, computers, etc.) that humans have developed. Its most important role right now might be in sharing and disseminating knowledge. History teaches us that people have always done this kind of abstract thinking we call mathematics. The future of math won't be that different in essence; it will just look different with new tools. For example, in the Middle Ages, making big tables of calculations was considered doing mathematics. Now we have calculators and computers that do that, so mathematicians work on other things. Similarly, AI might do some tasks we used to do, so mathematicians will explore new frontiers – maybe guided by AI, maybe by human insight. I believe humans won't stop doing mathematics or discovering new things. In fact, AI might show us things we never would have discovered on our own, which is exciting. We can look at the future of mathematics from that hopeful standpoint.

Number, Form, and Binary: A Short History of India's Mathematical Tradition and Its Relevance to the Global Pursuit of Artificial Intelligence

Saroj Kumar Rath and Abhishek Priyadarshi

University of Delhi

Adyan Rigsadi Rath

Ryan International School, Vasant Kunj, New Delhi

DOI: [10.1201/9781042018116-4](https://doi.org/10.1201/9781042018116-4)

ABSTRACT: This paper traces the evolution of India's mathematical imagination from the Harappan Era through the Vedic age to its enduring global influence, situating it within the larger history of human thought and the rise of artificial intelligence. Beginning with the Rigveda and Yajurveda, where number words such as *eka*, *daśa*, *śata*, and *sahasra* appear in liturgical contexts, it highlights how India developed a sophisticated vocabulary for powers of ten, culminating in terms like *parārdha*. The Śulva Sūtras (c. 800–200 BCE) extended this tradition into applied geometry, prescribing precise, algorithmic rules for altar construction and providing verbal formulations of the Pythagorean theorem and $\sqrt{2}$. The paper then follows the conceptual leap of *śūnya*—from a philosophical “void” in Pingala's binary prosody to a fully operational zero in Brahmagupta's *Brāhmasphuṭasiddhānta* (628 CE)—a breakthrough that transformed mathematics worldwide. Figures such as Āryabhaṭa, Bhāskara II, and the Kerala school demonstrate India's innovations in algebra, trigonometry, combinatorics, and algorithms. Through transmission to the Islamic world and Europe, these ideas laid the foundations for modern numeration and computation. By connecting Vedic geometry, linguistic numeration, and algorithmic thinking with today's digital logic, the paper argues that India's mathematical legacy is central to the global quest for artificial intelligence.

Keywords: Sulva Sutra, India, Greek, Numerals system, decimal system, Zero, Artificial Intelligence, Geometric proportion, Algorithm

1 METHODOLOGY

The methodology employed in this research manuscript adopts a triangulated framework, integrating historical-philological analysis, comparative historical inquiry, and algorithmic reconstruction. The philological component ensures rigorous fidelity to ancient Sanskrit texts,

capturing their precise mathematical lexicon and conceptual explanations. The comparative-historical approach resonates Indian mathematical innovations within the broader trajectory of global knowledge systems, tracing their transmission and influence across civilizations. Algorithmic reconstruction operationalizes these ancient procedural rules within contemporary computational logic, thereby establishing a substantive conceptual bridge to artificial intelligence. Collectively, this integrative methodology enables a robust demonstration that artificial intelligence is not merely a 20th-century phenomenon but represents the latest stage in a continuous, millennia-spanning tradition of systematic algorithmic reasoning.

2 KNOWING NUMBERS? A SHORT INTRODUCTION

The history of number terminology in ancient India was a very popular subject. Numerals in words in the Sanskrit language and in various scripts have been in vogue in ancient India for millennia. The Ṛgveda, which dates back to before 3000 BCE, is the earliest text where evidence of number words is found. Terms such as *eka* (one), *daśa* (ten), *śata* (hundred), and *śahasra* (thousand) appear frequently in hymns, often in cosmological or ritual contexts (Ṛgveda 1.164.10; [Macdonell 1916](#): p. 117; [Keith 1925](#): p. 211). Scholars have long debated the age of the Ṛgveda. Western thinkers, following the Aryan Invasion Theory, dated it between 1500 and 1100 BCE. Their estimate is not too far from the truth, yet it misses an important point. The Ṛgveda mentions the Saraswati River, which had vanished around 1900 BCE. This shows that the Ṛgveda must have been composed much earlier, when the Saraswati was still flowing (Shrikant, 2000, p. 107-152). The Rigveda devotes four verses to Saraswati in Mandala 2, Sukta 41, verses 15–18. These hymns vividly testify to the grandeur of the Rgvedic civilization and affirm its deep antiquity—an antiquity that cannot be confined within the narrow frame of 1500 BCE.

Around 1900 BCE, the Sutlej River shifted westward to join the Indus River, and the Yamuna River shifted eastward to join the Ganga River. This caused the Sarasvatī River to lose its perennial glacial sources and dry into a monsoon-fed seasonal stream. The drying of the Sarasvatī coincides with the decline of Harappan settlements in the Ghaggar-Hakra basin. The Harappan phase shifted eastward towards the Ganga-Yamuna doab after 1900 BCE. (Pal, Sahai, Sood and Agrawal, p. 317).

In other words, Indians were well adept at calculation, including addition, subtraction, and the use of zero as a concept, although they did not initially employ written numerical symbols. Instead, they expressed numbers through words in Sanskrit or encoded poetically in verses. A notation for powers of ten up to 10^{17} was already in use from Vedic times, with single words employed to denote each power.

Found in the Ṛgveda, with 16 hymns, the ‘Purusasukta’ or ‘the Hymn of the Cosmic Being’ talks about numbers. ‘The cosmic person has a thousand heads, a thousand eyes, and a thousand feet. Encompassing the whole earth, he extends beyond it by ten fingers’ breadth’ (Ṛgveda, Maṇḍala 10, Sūkta 90)¹. In this hymn, Sanskrit term ‘*śahasra*’ denotes ‘a thousand,’ while the word ‘*atyatiṣṭha*’ conveys the mathematical sense of ‘what remains over’ or the remainder after subtraction. Similarly, the term ‘*aṅgulaṃ*’ is better understood as a unit of length that is finger length rather than as ‘fingers.’

In the tenth *maṇḍala* of the Ṛgveda², terms such as *pramā* (‘measure’), *paridhiḥ* (‘circumference’), and *prāṅga* (‘triangle’) suggest an early concern with geometric and

mathematical ideas. The notion of dividing the circle into 360 degrees may well have its roots in this Vedic worldview. One hymn declares: *‘The twelve are the months, the wheel is but one; it has three hubs—who can comprehend it? Within it are set three hundred and sixty spokes, both fixed and in motion³.’* This imagery reveals a strikingly sophisticated conception of cyclic order, time, and measure in the earliest stratum of Indian thought.

Numerous arithmetic terms occur in the R̥gveda, and numerical figures and values had already been established during the Vedic period. The following table presents the values of various figures current in R̥gvedic times.

Table 1. Vedic numbers in word characters.

R̥gvedic term (in Sanskrit)	Value	Source
एकं संवत्सरम्	One year	R̥gveda, Mandala 1, Sukta 110, Sloka 4
चतुर्भिः साकं नवतिम्	Ninety-four	R̥gveda, Mandala 1, Sukta 155, Sloka 6
नवभिः साकं नवतिम्	Ninety-nine	R̥gveda, Mandala 1, Sukta 54, Sloka 6; R̥gveda, Mandala 1, Sukta 84, Sloka 13
त्रीणि शता त्री सहस्राणि त्रिंशच्च देवाः	3309 Gods	R̥gveda, Mandala 3, Sukta 9, Sloka 9
सहस्रा त्रिंशत्	30,000	R̥gveda, Mandala 4, Sukta 30, Sloka 21
शतं सहस्रम्	One lakh	R̥gveda, Mandala 2, Sukta 14, Sloka 6
शतमा सहस्रम्	Many lakhs	R̥gveda, Mandala 2, Sukta 14, Sloka 7

In Purusasukta hymns of the R̥gveda⁴ it was analyzed how the subtraction of a quarter from one is given as three-quarters. Mathematically, it is correct as explained below.

$$1 - 1/4 = 4/4 - 1/4 = 3/4, \text{ so } 1 - 1/4 = 3/4.$$

Scholars of the history of mathematics note that this is one of the earliest literary references to fractional division ($\frac{1}{4}$ and $\frac{3}{4}$) and a cosmological application of subtraction and partitioning in the Vedic corpus. The Krishna Yajurveda contains references to several fractional divisions: *kalā*, denoting $\frac{1}{16}$, *kuṣṭha*, $\frac{1}{12}$, *śapha* $\frac{1}{8}$, and *pāda*, $\frac{1}{4}$ ⁵. These fractional terms are mentioned in the context of bargaining over the price of a commodity—in this case, Soma—to be purchased. In later colloquial usage, similar fractional expressions survived in currency terms—for example, *Chār ānnā* meaning one-quarter of a rupee (25 paise), *Ath ānnā* one-half of a rupee (50 paise), *Bārah ānnā* one-third of a rupee, and *Solah ānnā* signifying a full rupee (100 paise).

In the Yajurveda, a systematic series of odd and even numbers is mentioned. The odd numbers include 1, 3, 5 ... up to 31 and 33, while the even numbers extend as 4, 8, 12 ... up to 48⁶. The series of fractions is given as 1, $1\frac{1}{2}$, 2, $2\frac{1}{2}$, 3, $3\frac{1}{2}$, 4.

These scriptures demonstrate that the decimal system was firmly embedded in Indian culture as early as the third millennium BCE. Texts such as the *Yajurveda* in its ritual descriptions, and the *Mahābhārata* and *Rāmāyaṇa* in their accounts of statistics and measurements, make frequent and uninhibited use of these terms. (Struik, 1948, 38).

The Shukla Yajurveda describes the use of large numbers as well. The sequence was first reported in the Shukla Yajurveda, which runs as follows: *eka*, *daśa*, *śata*, *sahasra*, *ayuta*, *lakṣa*, *prayuta*, *koṭi*, *arbuda*, *abja*, *kharva*, *nikharva*, *mahāpadma*, *śaṅkha*, *jaladhi*, *antya*, *mahāśaṅkha*, and *parārdha*. In a celebrated hymn⁷ the long sequence of numbers is recited in a litany, culminating in very large powers of ten.

Table 2 Table 2 Vedic numerals in word without symbol.

Sanskrit Term	Place Value	Equivalent in Powers of Ten
(एक) Eka	1	
(दश) Daśa	10	
(शत) Śata	100	
(सहस्र) Sahasra	1,000	
(अयुत) Ayuta	10,000	

Sanskrit Term	Place Value	Equivalent in Powers of Ten
(लक्ष) Lakṣa	100,000	
(कोटि) Koṭi	10,000,000	10 million
(पद्म) Padma	10 ¹⁴	100 trillion
(परार्ध) Parārdha	10 ¹⁷	100 quadrillion
(महाशङ्ख) Mahāśaṅkha	10 ²¹	Sextillion
(वृन्दा) Vrindā	10 ²³	100 sextillion
(खर्व) Kharva	10 ²⁷	Octillion
(निकर्व) Nikharva	10 ²⁹	100 octillion
(महापद्म) Mahāpadma	10 ³³	Decillion
(शङ्ख) Śaṅkha	10 ³⁷	100 decillion
(जलधि) Jaladhi	10 ⁴¹	Duodecillion

Sanskrit Term	Place Value	Equivalent in Powers of Ten
(अन्त्य) Antya	10^{45}	Quattuordecillion
(मध्य) Madhya Paraardha of Śaṅkha	10^{49}	100 quattuordecillion
	And so on ...	Infinity in theory

After the Vedas, the Brāhmaṇas and Upaniṣads refined such enumerations, employing them systematically in sacrificial calculations, numerical classifications, and metaphysical speculations about the universe (Śatapatha Brāhmaṇa 10.4.2.23; Eggeling, 1882, p. 89; Olivelle, 1996, p. 54). By the period of the Mahābhārata and the Purāṇas, the lexicon of numbers had expanded dramatically. Larger terms such as *lakṣa* (100,000), *koṭi* (ten million), *padma*, *śaṅkha*, *kharva*, *nikharva*, *mahāpadma*, *jaladhi*, and *parārdha* were employed, particularly in the description of vast time cycles (yugas, kalpas, manvantaras) and cosmological magnitudes⁸ (Mahābhārata, Śānti Parva, Book 12, Chapter 231; Hopkins, 1915, p. 155; Wilson, 1840, p. 31).

A striking example comes from the eldest Pāṇḍava, Yudhiṣṭhira, when he declared in the Kuru dynasty's assembly that 'When opposed to one another, we are five and they are a hundred; yet when facing an external foe, we are a hundred and five together'⁹

(Vyasa, Mahābhārata, Adhyaya 26, Verse 7). This reveals that Yudhiṣṭhira not only understood the numbers five, one hundred, and one hundred and five, but also expressed them in words rather than symbols. Not only this, the famous dice game episode in the same chapter of the Mahābhārata reveals an understanding of chance and numerical play. While not probability theory in the modern sense, the repeated emphasis on throws of numbers, stakes, and odds shows an intuitive grasp of numerical randomness. In that age, knowledge was preserved orally, and every numerical value was conveyed through spoken composition.

In this series of Indian knowledge traditions, the Jaina canon went further. Works such as the *Anuyogadvāra-sūtra* and the *Tiloyapaṇṇatti* from circa 3rd–5th century CE provided systematic classifications of extraordinarily large numbers, often extending beyond practical calculation into the realm of philosophical infinity. From *Anuyogadvāra-sūtra*, verse 71, as found in a Karmagrantha commentary, 'There is one type of saṅkhyāta (countable). There are three types of asaṅkhyāta (innumerable), and three types of ananta (infinite). Thus, there are a total of seven types of numbers. Each type has three sub-types: jayamyā (inferior), madhyāmā (intermediate), and utkrṣṭa (superior)' (Tiloyapaṇṇatti 2.25; Jaini, 1979, p. 92; Jain, 2011, p. 144; [Datta and Singh, 1935](#), p. 22). Thus, long before the invention and adoption of written numeral symbols, India had cultivated an advanced tradition of numerical thought expressed through linguistic formulation, showing that the intellectual framework for mathematics preceded its symbolic notation.

Historian Will Durant wrote admiringly about India's role in the development of mathematics in his monumental series *The Story of Civilization*, where he explicitly credits India with the invention of numerals and the decimal system, including zero. He wrote,

‘India was the mother of our philosophy, of much of our mathematics, of the ideals embodied in Christianity ... She was the mother of our numeral system, of the decimal system, of the zero. Without her discoveries no worthwhile mathematics would have been possible’ ([Durant, 1935](#), p. 527).

The number zero is the subtle yet profound gift of ancient India to humanity. This concept stands among the most significant inventions in the ascent of man, shaping the growth of culture and civilization. The Indian called it *śūnya*, a term drawn from philosophical literature, where it signified the ‘void’.

3 PRACTICAL APPLICATION OF MATHEMATICAL WISDOM

The Śulva Sūtras are among the earliest mathematical texts of India, and they are indeed prime evidence of advanced mathematical understanding in the Vedic period. Explaining the Śulva Sūtras is a task of exceptional scholarly sensitivity, as these works are deeply rooted in a long and intricate development of the Sanskrit literary tradition. Veda, especially Rg Veda, is at the top of this ladder. This progression begins with the foundational authority of the Vedas, is expanded and systematised through the Vedāṅgas, and ultimately takes form in the concise, aphoristic style that defines the Sūtra literature. The Śulva Sūtras cannot be meaningfully understood in isolation. They represent a specialised branch of the Kalpa Vedāṅga, whose ritual and mathematical prescriptions are inextricably linked to the broader explanatory, linguistic, and ritual frameworks developed over centuries. Their content reflects not merely technical instructions for altar construction, but the cumulative intellectual inheritance of Vedic ritualism—preserved through oral transmission, codified in precise formulae, and situated within a tradition that privileges both fidelity to inherited authority and the refinement of specialised knowledge. Consequently, any examination of the Śulva Sūtras must navigate a layered textual heritage, one that demands simultaneous attention to philological precision, historical context, and the internal logic of Vedic ritual science ([Thibaut, 1875](#), p.1-5).

The Śulva Sūtras are a group of four extant ancient Indian texts on applied geometry, composed as appendices to the Śrauta Sūtras within the Vedāṅga group of texts. The six Vedāṅgas (‘limbs of the Veda’) are auxiliary disciplines developed to ensure the correct preservation, interpretation, and performance of the Vedic texts¹⁰. They emerged between the late Vedic and early post-Vedic periods as systematic sciences supporting the oral tradition. These are: Śikṣā (phonetics), which codifies correct pronunciation and accentuation; Chandas (prosody), concerned with the metrical structure of Vedic hymns; Vyākaraṇa (grammar), which analyses linguistic form to safeguard textual integrity; Nirukta (etymology), providing semantic explanations of obscure Vedic words; Jyotiṣa (astronomy/astrology), prescribing the calculation of ritual calendars; and Kalpa (ritual science), which sets out procedural rules for sacrificial and domestic rites, including the geometrical prescriptions of the Śulva Sūtras. Collectively, the Vedāṅgas reflect the intellectual and technical infrastructure that ensures the precise transmission and ritual enactment of Vedic knowledge (Ghosh, 1938, xix).

Among the Vedangas, the Kalpa Śāstra—not to be confused with the *Kalpa Sūtra* of the Jain tradition—is the root source of Sulva Sutra ([Pāṇini, 1991](#), p. 556). In the Vedic framework, *Kalpa* denotes the codified science of ritual procedure, encompassing the precise methodologies for conducting sacrificial ceremonies (*yajña*), domestic rites, and other religious observances, while

the term *śāstra* signifies a systematic treatise. As such, the Kalpa Śāstra constitutes a mass of theoretical and practical knowledge governing Vedic ritual law and praxis. Within this tradition, four major textual subdivisions emerged: the Śrauta Sūtras, prescribing the execution of large-scale public sacrifices grounded in the *śruti* (revealed) canon. The Gṛhya Sūtras, detailing domestic rituals such as marriage, naming, and funerary rites; the Dharma Sūtras, addressing normative moral, legal, and socio-religious codes; and the Śulva Sūtras, appended to certain Śrauta Sūtras, devoted specifically to the geometric and mathematical principles of altar construction ([Sarup, 1920](#), p. 18).

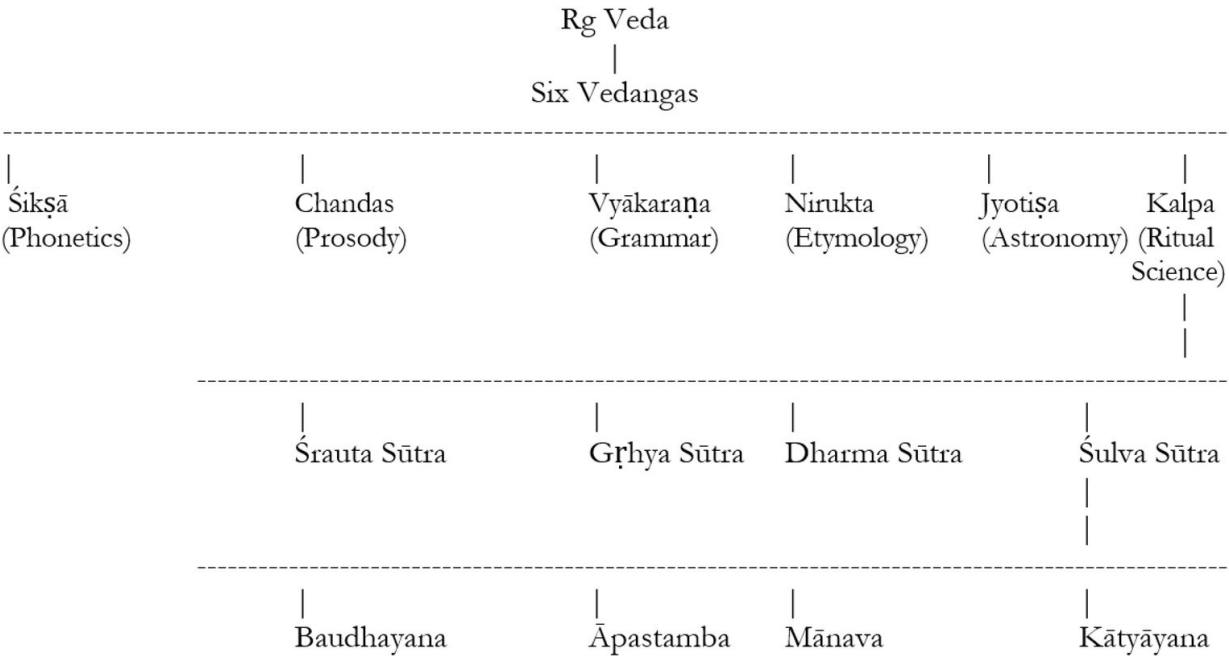


Figure 1. An explanation about sulva texts. Source: Prepared by the authors in consultations with the original Vedic texts.

Composed between the 8th and 2nd centuries BCE, the Śulva Sūtras represent the earliest extant Indian treatises on applied geometry, embedded within Vedic ritual practice ([Thibaut, 1875](#), 23). The primary purpose of the Śulva Sūtras was to provide precise, algorithmic instructions for constructing Vedic fire-altars (vedi/chithi) of prescribed shapes and dimensions, using cord-based measurement techniques. The term *śulva*¹¹ ([Thibaut, 1875](#), p.1) means ‘measuring cord’ and *sūtra*¹² (Kielhorn, 1880, p.4) means ‘aphoristic rule,’ reflecting their practical, procedural nature.

Only four complete Śulva Sūtras are extant, each attributed to a prominent ritual authority belonging to a distinct Vedic school. The *Baudhayana Śulva Sūtra*, attributed to Baudhayana and dated to approximately 800–500 BCE, is generally regarded as the earliest and most influential, laying the foundational principles of altar geometry. The *Āpastamba Śulva Sūtra*, ascribed to Āpastamba (c. 600–300 BCE), introduces further refinements in geometric procedures and construction methods. The *Mānava Śulva Sūtra*, attributed to Mānava (c. 500–300 BCE), reflects distinctive ritual prescriptions and structural designs specific to its school's tradition. The *Kātyāyana Śulva Sūtra*, credited to Kātyāyana (c. 300–200 BCE), is considered the most recent, presenting a more systematised and consolidated approach to geometric instruction ([Srinivasachar](#)

[and Narasimhachar, 1931](#), p.28). In all cases, the attributed author was likely the head of a Vedic ritual lineage, and the composition represents an oral tradition later committed to writing, serving as a codification of that school's precise mathematical and geometric practices in the service of sacrificial rites. What is of particular significance is the systematic codification of mathematical formulae and geometric proportions governing shape and size, accompanied by a sustained tradition of refinement in the sciences of mathematics and geometry, whereby empirical techniques were progressively formalised into precise, reproducible rules, ensuring both the preservation and continuous development of technical knowledge across generations.

Not only the *Sulva Sutra* but also archaeological excavations, especially at Kalibangan, have significantly enriched scholarly debates on the antiquity of mathematical knowledge and the scientific basis of Vedic ritual traditions as codified in the *Kalpa Śāstra*. Archaeological findings reveal the presence of a series of meticulously constructed fire altars within two citadel areas at Kalibangan, located deep within Rajasthan ([Thapar, 1975](#), p.27). These structures display a striking adherence to precise geometrical configurations and proportional systems, strongly reminiscent of the ritual altar designs later formalised in the *Śulva Sūtras*.



Figure 2. Fire Altars in Indus Valley. Source: B.K. Thapar, *Kalibangan: A Harappan Metropolis Beyond the Indus Valley*, Expedition Magazine, Vol. 17, No.2, Winter, 1975, pp. 25.

What renders this finding particularly remarkable is the chronological disparity between the two traditions. The codification of fire altar formulae in the *Śulva Sūtras*—textual appendices to the

Śrauta Sūtras—can be securely placed within the period c. 800 BCE onwards. In contrast, the Indus Valley Civilization, to which Kalibangan belongs, flourished much earlier, between approximately 3500 and 2500 BCE ([Lal, 2002](#), p.115-120). The evidence suggests that the advanced geometrical principles and mathematical formulae, later articulated in the Śulva Sūtras, may have had antecedents in the urban and ritual culture of the Harappan world. This continuity points toward a deep-rooted indigenous tradition of mathematical and geometric reasoning, integrated into both civic and ritual architecture, and challenges the conventional view that such knowledge systems were exclusively products of the later Vedic age.

4 SHAPES AND SYMBOLISM OF FIRE-ALTARS IN GEOMETRIC CONTEXT

In contrast to the monumental but socially distant mathematical achievements of ancient Egypt, Greece, Mesopotamia, and Mesoamerica, the Śulva Sūtras testify to a unique Indian integration of advanced numerical calculation, precise measurement, and algorithmic geometry into the fabric of daily life as well as collective religious practice ([Rath, 2025](#), p.178). Indians codified indispensable procedural knowledge for altar construction. For example, the *Baudhayana Śulva Sūtra* warns that even the smallest deviation in prescribed measurements could compromise the very efficacy of the sacrificial rite (Dvivedi, 1890, p.15).

What distinguished the Indian civilisation from its Egyptian, Mesopotamian, and Hellenic counterparts was the systematic integration of advanced mathematical knowledge into the structures of daily life ([Rath, 2025](#), p.178). In the Indian context, disciplines such as arithmetic, geometry, and algorithmic computation were not confined to monumental architecture or elite scholarly pursuits, but were embedded within routine cultural, religious, and domestic practices—most notably exemplified by the Śulva Sūtras. Archaeological excavations ([Thapar, 1975](#), p. 27) at Kalibangan noted that, in addition to the series of fire altars discovered within the citadel complex, certain domestic structures also contained what appeared to be fire altars integrated into the interior layout of individual rooms. These domestic altars, often constructed on a smaller scale than their ceremonial counterparts in public or communal spaces, were typically composed of well-prepared mud-brick platforms with clearly defined shapes—frequently circular or rectangular—and showed evidence of repeated burning.

The presence of such installations within private dwellings suggests that ritual fire worship was not confined solely to large-scale, state-organised ceremonies, but also formed part of the household religious sphere. This finding provides important archaeological corroboration for later Vedic textual traditions, such as the *Grhya Sūtras*, which prescribe the maintenance of sacred fires within the home for daily domestic rites. In this respect, Kalibangan's domestic fire altars represent an early material antecedent of household ritualism, thereby underscoring the continuity of ritual fire cults from the Harappan period into the later Vedic age. Such evidence strengthens the argument for a long-standing cultural embedding of geometry, measurement, and ritual space design in both public and private domains of ancient Indian society ([Allchin and Allchin, 1982](#), p.227).

The Śulva Sūtras prescribe diverse altar shapes—squares, rectangles, circles, trapezia, and animal forms such as falcon-chithi, tortoise-chithi, and rhomboids—each symbolic of a specific ritual goal. For instance, a falcon-shaped altar signified attainment of heaven, a tortoise shape invoked Brahmanic stability, while rhomboid shapes were for subjugating enemies (Dvivedi, 1890, p.7).

Moreover, the manipulation of areas—transforming one shape into another while preserving area—became a hallmark of Śulva geometry. This includes methods to transform rectangles to squares, construct circles equal in area to squares (“squaring the circle”), and vice versa.

5 GEOMETRIC TECHNIQUES AND CALCULATIONS

All four principal Śulva Sūtras state the Pythagorean theorem verbally, centuries before Greek geometrician Pythagoras himself (c. 570–495 BCE). Baudhayana expresses it succinctly: “The diagonal rope of a rectangle produces both [i.e., the areas] which the two sides produce separately” (Buhler, 1882, 67)¹³.

These sūtras also provide Pythagorean triples—integer sets like (3, 4, 5), (5, 12, 13), (8, 15, 17)—derived to meet altar-construction needs (Dani, 2003, 221). The impetus was practical: rituals frequently demanded altars of equal area but differing shape, necessitating integer-dimensioned right triangles to facilitate area-equivalence while preserving whole-brick counts.

Baudhayana also delivers a remarkably accurate approximation to $\sqrt{2}$ in a single verse: “Increases by its third, and that third by its own fourth, less the thirty-fourth part of that fourth ...”, yielding 1.414215 ..., correct to five decimal places (Dvivedi, 1890, p.66)¹⁴. This degree of precision, achieved via an algorithmic sequence of fractional operations, illustrates an implicit understanding of irrational numbers.

The construction of Vedic fire altars, as prescribed in the Śulva Sūtras, demanded a level of mathematical precision that extended far beyond elementary measurement, necessitating sophisticated computational methods. These ritual geometries often required the transformation of one shape into another of equivalent area (e.g., converting a square altar into a circular or falcon-shaped one), the scaling of dimensions according to complex proportional rules, and the iterative approximation of irrational quantities such as $\sqrt{2}$.

Such operations, particularly when involving large or fractional dimensions, would have been greatly facilitated by the conceptual apparatus of the place-value numerical system and the symbolic representation of zero—innovations that, while attaining their most explicit formulation in later Indian mathematics, were rooted in the computational culture that the Śulva Sūtras exemplify (Seidenberg, 1962, p.488-527). The capacity to represent large numbers succinctly, perform multi-step arithmetic, and track intermediate results without ambiguity would have been indispensable in designing altars whose dimensions adhered to exact geometric ratios.

In practice, these calculations ensured that ritual prescriptions were satisfied not merely in form but in precise quantitative detail. For instance, the *Baudhayana Śulva Sūtra*'s rules for altar enlargement across successive years of a soma sacrifice required proportional increases in area while maintaining strict geometric integrity—a process that, without a coherent numerical framework, would risk cumulative error and thereby undermine ritual efficacy (Buhler, 1897, p.213). This intersection of symbolic mathematics, geometric construction, and ritual cosmology represents a distinctive hallmark of ancient Indian intellectual culture, in which mathematical thought was not an abstract pursuit but an integral component of religious life, encoded and perpetuated through the authoritative medium of the Kalpa Śāstra tradition.

The sūtras offer explicit geometric recipes for area-preserving transformations, for instance, transforming a rectangle into a square by “cutting off a square, halving the remainder, and reattaching it”. They also treat circle–square equivalences: Apastamba notes implicitly that the

parts of the square outside the inscribed circle equal those inside, hinting at a trade-off based on rope length and intuitive proportionality ([Datta and Singh, 1935](#), p.22).

The value of π has long fascinated civilizations, and much of the geometry in the Śulva Sūtras implicitly revolves around its application in altar construction and circle-square transformations ([Sarasvati, 1982](#), p. 52). What is striking, however, is that Pythagoras, who was present during circa 570 and 495 BCE, and the early Greek mathematical tradition did not engage in the direct computation of π . Pythagoras is remembered primarily for the Pythagorean theorem and for his philosophical emphasis on number mysticism, integer ratios, and harmonic relations, rather than for work on circular constants¹⁵ ([Heath, 1921](#), p.73).

By contrast, the earliest textual evidence of mathematical reflection on π is found in ancient Egypt, notably in the Rhind (Ahmes) Papyrus (c. 1650 BCE), where geometric procedures yield approximations of the circle's area (Peet, 1923, p.64). Meanwhile, archaeological discoveries at Kalibangan, a site of the Indus Valley Civilization (c. 3500–2500 BCE), indicate the use of geometry in the construction of fire altars, though such practices remained embedded in ritual architecture. The formal codification of these procedures emerged much later in India, with the Śulva Sūtras (c. 800 BCE onward), where altar geometry and mathematical reasoning were systematically recorded in a concise aphoristic style.

Within these texts, the value of π was empirically approximated through altar-construction needs, with estimates ranging from 2.99 to 3.2022 depending on context ([Datta and Singh, 1962](#), p.21). This demonstrates an iterative, practical approach: rather than defining π abstractly, the Vedic ritualists worked out sufficiently precise approximations to ensure ritual efficacy. Such calculations reveal both the scientific acuity and ritual precision of early Indian mathematics, predating by centuries ([Kline, 1972](#), p.41) the more systematic treatment of π by Archimedes of Syracuse (c. 287–212 BCE).

The Rhind Mathematical Papyrus, also known as the Ahmes Papyrus from Egypt, is one of the intriguing documents on the development of the value of π . Much before the Sulva Sutras were compiled, the Egyptians had documented mathematical tradition in their peculiar way. The document was copied by the scribe Ahmes around 1650 BCE, during Egypt's Second Intermediate Period, though Ahmes himself stated that the work was based on older sources dating back to the Middle Kingdom c. 2000 BCE (Peet, 1923, p.4). The papyrus was purchased in 1858 at Luxor by the Scottish antiquarian Alexander Henry Rhind, whose name it now bears, and is preserved today in the British Museum¹⁶.

We don't know who the original author of the Rhind Papyrus was. While Ahmes' copy is the only extant version preserved today, the true authorship of the pioneering work remains unknown. The papyrus, therefore, provides a unique but indirect window into a mathematical tradition that predates its surviving manuscript by several centuries.

The Ahmes Papyrus is essentially a mathematical handbook containing 84 problems with solutions, written in hieratic script. Its content ranges from simple arithmetic operations, unit fraction conversions, and proportions to more advanced problems involving linear equations, geometric progressions, areas, and volumes. Particularly notable is its treatment of circular areas, where Ahmes equates the area of a circle to that of a square whose side equals $\frac{8}{9}$ of the diameter. This procedure implies a value of $\pi \approx 3.1605$, remarkably close to the true value, and testifies to the Egyptians' practical ingenuity in geometry (Peet, 1923, p.25).

While the Ahmes Papyrus does not pursue abstract theory or axiomatic reasoning, it reflects the applied and utilitarian nature of Egyptian mathematics, closely tied to land measurement, construction, and administrative tasks ([Gillings, 1971](#), p.520). Thus, it represents an important

stage in the global history of mathematics, situated between the pragmatic problem-solving of early civilizations and the systematic deductive traditions of later Indian and Greek mathematics.

The Rhind Papyrus illustrates the empirical-practical character of Egyptian mathematics, oriented toward administrative, architectural, and surveying needs rather than deductive proof. In this sense, it stands in methodological contrast both to the Śulva Sūtras of India, which encoded geometry within ritual practice, and to the Greek axiomatic tradition, where Archimedes and Euclid systematised mathematics into logically ordered demonstrations.

In contrast to the axiomatic approach of Greek geometrician Euclid, Śulva geometry remains algorithmic and instrumental: instructions are stepwise without formal proof, but many implicitly rely on geometric reasoning, preserving rigour through construction protocols (Sneidenberg, 1962, p.488).

Euclid's (c. 300 BCE) seminal work, the *Elements*, organised and codified the mathematical knowledge of his time into a systematic and logical framework. Little is known about his life, though Proclus Lycaeus records that he taught in Alexandria during the reign of Ptolemy I during 323–283 BCE (Morrow, 1970, p. 24). Proclus Lycaeus, who lived during circa 412–485 CE, was a late Neoplatonist philosopher and mathematician, and one of the last great commentators on Greek philosophy and science. Born in Constantinople and educated in Alexandria and Athens, he became the head (*diadochos*) of the Platonic Academy in Athens.

The axiomatic nature of Euclid's geometry refers to the systematic and deductive way in which Euclid, in his *Elements*, built the entire structure of geometry from a small set of self-evident assumptions. Instead of relying on empirical observation or ad hoc rules, Euclid established a rigorous logical framework, where all geometric truths were derived step by step from foundational principles. Euclid began by defining basic terms (point, line, plane, etc.), and then laid down axioms (common notions) and postulates (geometric assumptions). For example, the statement “a straight line can be drawn joining any two points” was taken as universally true, not requiring proof (Heath, 1908a, p.117).

But did Euclid talk about pi? The answer is negative, but although in his *Elements*, he does not explicitly discuss π as a constant in the way later mathematicians like Archimedes did, Euclid proved several geometrical propositions about circles that implicitly involve π . But he never attempted to calculate or approximate its value. For example, in *Elements*¹⁷ Euclid shows that the areas of circles are proportional to the squares of their diameters (Heath, 1908b, p.381). Symbolically, if two circles have diameters d_1 and d_2 , then their areas satisfy:

$$\frac{A_1}{A_2} = \left(\frac{d_1}{d_2} \right)^2$$

Figure 3. Formulating Pi.

This relation assumes a constant ratio between the circumference (or area) of a circle and its diameter, which is precisely what we today call π . However, Euclid stopped short of identifying or computing that constant. Thus, while Euclid laid the axiomatic foundation that made the definition of π possible, it was Archimedes of Syracuse who first treated π as a mathematical problem in its own right, providing a rigorous numerical approximation (Heath, 1908a, p.22).

The next significant advance in the historical treatment of π came from Archimedes of Syracuse, who stands as one of the earliest mathematicians to subject the constant to a rigorous geometric investigation. Employing the method of exhaustion—a precursor to integral calculus—Archimedes approached the problem of determining the ratio of a circle's circumference to its diameter by systematically inscribing and circumscribing regular polygons within and around a circle. By increasing the number of polygonal sides to as many as 96, he was able to establish numerical bounds for π with remarkable precision for his era. Archimedes concluded that π must lie between 3.1408 and 3.1429 (Heath, 1897, Prop. 3), thereby providing a rational approximation that would dominate mathematical thought for centuries.

What distinguishes Archimedes' approach from earlier empirical traditions, such as those preserved in the Śulva Sūtras, is his methodological rigour. Instead of relying on practical rules derived from measurement, he developed a logically consistent framework rooted in deductive geometry. In this sense, Archimedes represents the transition from practical approximation to formal proof, laying the foundations for the systematic study of mathematical constants in the Greco-Hellenistic world (Kline, 1972, p.57).

The Śulva Sūtras developed geometry practically, as rules and constructions to ensure correct proportions of Vedic fire altars, embedding mathematics in ritual and daily religious practice. Euclid, by contrast, approached geometry theoretically, building an abstract, axiomatic-deductive system independent of ritual or utility, aimed at universal logical consistency. Śulva geometry is ritual-oriented and pragmatic; Euclidean geometry is abstract, systematic, and axiomatic.

Within the Śulva texts, different approximations of π appear, reflecting both an empirical and iterative approach. For example, some texts effectively employ $\pi \approx 3.2022$ (by approximating the circumference of a circle as three times its diameter plus a slight adjustment), while others come closer to $\pi \approx 2.99$, depending on the particular transformation of shapes being described (Datta and Singh, 1962, p.34). These values indicate that the Vedic scholars were not attempting to define π as a universal constant in the abstract Euclidean sense, but rather to achieve a functional accuracy suited to the precision required in ritual practice.

What is striking is the contextual flexibility. Instead of prescribing a single universal value, the Śulva Sūtras allowed different working approximations, each derived through practical geometric constructions involving ropes, diagonals, and proportional measures. This empirical method reflects a scientific mindset rooted in observation, adaptation, and iterative refinement, where the success of the ritual outcome validated the accuracy of the computation (Seidenberg, 1962, p.488).

Thus, the Śulva approximations to π predate and conceptually differ from the later Greek pursuit of mathematical constants through abstract deduction. While Euclidean mathematics eventually framed π as part of a formal axiomatic system, the Vedic mathematicians approached it as a ritually embedded, practically tested constant, demonstrating the integration of mathematical reasoning into the fabric of cultural-religious life (Datta and Singh, 1962, p.37).

The Sulva texts employ visual geometry, cord constructions, and arithmetic. Users manipulated rajju (cord ropes) as compasses, straight-edges, or measuring lines to lay out the altar footprint on the ground. Figures were built full-scale using stakes and cords, scaled for bricks and layers of fire-altars that could extend several metres.

Śulva texts influenced subsequent algebra and geometry in India. The ritual tradition persisted in schools and oral lineages; archaeological examples of altar shapes hint at continuity—though not conformity to prescribed designs. Modern scholars regard the Śulva Sūtras as an early step in Indian mathematical sophistication ([Thibaut, 1875](#), p.2). They present verbal proofs—concise syllogistic logic rather than Euclid's postulates—alongside algorithmic methods for squaring the circle, deriving irrational numbers, and manipulating shapes. Their geometry parallels Euclid in methodical construction while lacking formal axiomatic foundations.

In a contemporary pedagogical context, Śulva geometry offers an engaging example of applied mathematics, illustrating how abstract mathematical concepts can arise from concrete ritual tasks. The *Śulva Sūtras* codified geometric principles, including a remarkably accurate approximation of $\sqrt{2}$ and methods for transforming geometric shapes—for example, constructing a square with the same area as a circle. These were not merely theoretical exercises but practical, algorithmic procedures: step-by-step rules that anticipate what we now recognize as algorithms, the foundational building blocks of computer science and artificial intelligence ([Hayashi, 2003](#), p.367).

Ancient Egyptian mathematics, as evidenced in the *Rhind Papyrus*, emphasized practical arithmetic and geometry, primarily for construction, land measurement, and astronomy. Their work introduced the concept of encoding knowledge into repeatable procedures—a principle that lies at the heart of modern programming logic ([Gillings, 1972](#), p.9, 177 & 230). Greek mathematicians later formalized and expanded upon these ideas. Euclid's *Elements* laid the foundation for axiomatic geometry and systematic reasoning, while Archimedes' method of exhaustion anticipated the development of calculus. Although known primarily as a philosopher, Aristotle (384–322 BCE) developed foundational principles of logic and reasoning that have become essential to both modern mathematics and the theoretical bedrocks of machine learning ([Aristotle, 1928](#), p. 24).

6 THE QUEST TO DEFINE ZERO, EMPTINESS, VOID, SUNYA

Indian mathematicians, across the ages, pursued innovative approaches to mathematics. For instance, Pingala in 200 BCE, introduced the concept of binary numbers—departing from conventional calculations—an idea that later became the cornerstone of modern computer coding and programming. In his *Chandaḥśāstra* (treatise on prosody), Pingala dealt with describing and classifying Sanskrit poetic meters. He provided recursive, rule-based methods for counting syllable patterns. This is essentially an algorithmic system, remarkably similar to modern binary computation (Shastri, 1874, p.54).

To analyze meters, Pingala developed binary representations:

Laghu (short syllable) = 0

Guru (long syllable) = 1 (Shastri, 1874, p.84)

This is, in effect, the earliest known use of a binary system in world history—predating Gottfried Wilhelm Leibniz (1646–1716), the German philosopher and mathematician, who is credited with formally developing the binary numeral system nearly 1900 years later. In 1703, Leibniz published *Explication de l'Arithmétique Binaire*, explaining how numbers can be represented using only 1 and 0. Pingala described recursive procedures to generate all possible

combinations of laghu and guru for meters of any given length. These rules are algorithmic, which define base cases, apply step-by-step construction rules and enumerate all possible patterns systematically ([Sen and Bag, 1983](#), p. 27).

In Pingala's system, when describing combinations, he sometimes required a placeholder to mark absence in a metrical position. This placeholder was called śūnya (empty, void). While this is not yet the numerical zero of Brahmagupta, which he formalized in the 7th century, it is a conceptual precursor: using 'emptiness' as a meaningful marker in a formal system. Over time, this philosophical-mathematical notion of śūnya matured into the symbol 0, first seen in the *Bakhshali* manuscript ([Datta and Singh, 1935](#), p.214), which was fully formalized by Brahmagupta. What began as a method to catalogue poetic rhythms evolved into a powerful mathematical framework.

The Bakhshali manuscript found in Bakhshali village near Peshawar dates back to c. 3rd to 7th century CE (carbon dating shows layered composition). The script used in the manuscript is the Śāradā script, while the language is Sanskrit, and it was written in Birch bark. It was discovered in 1881 by a farmer, and the 70-page Birch-leaf manuscript has resided in the Bodleian Library, University of Oxford, since 1902. It uses a dot (·) as the zero placeholder in a decimal place-value system — one of the earliest known examples of zero's usage in such a mathematical framework ([Datta and Singh, 1935](#), p.218).

However, the dot symbol in the Bakhshali script is the one that ultimately evolved into the hollow-centred version of the symbol that we use today. It also sowed the seed for zero as a number, which is first described in a text called Brahmasphutasiddhanta, written by the Indian astronomer and mathematician Brahmagupta in 628 AD ([Colebrooke, 1817](#), p. 346)¹⁸. In a recent study, three folios of the Bakhshali Manuscript were radiocarbon-dated at the Oxford Radiocarbon Accelerator Unit ([Devlin, 2017](#)). The results revealed that the samples originate from three distinct centuries: one from 224–383 CE, another from 680–779 CE, and the third from 885–993 CE—raising important questions about how materials spanning such a wide temporal range came to be compiled as a single manuscript. Was it a student's learning collection, a teacher's compiled notes or an institutional record (e.g., monastery or gurukul) or a copying over of older texts onto newer material? We never know.

But what does the manuscript talk about? Written in a form of Sanskrit, translations suggest it served as a training manual for merchants trading along the Silk Road, containing practical arithmetic exercises and early forms of algebra ([Lavis, 2024](#), p.85-92 & p.380-386).



Figure 4. Earliest Trace of Zero. In this close-up image, a dot appears as a placeholder in the bottom line—a precursor to the modern numeral zero. Photograph: Courtesy of MS. Laud Misc. 671, Bodleian Libraries, University of Oxford.



Figure 5. Original Bakhshali Manuscript. Bakhshali Manuscript, Courtesy of MS. Laud Misc. 671, Bodleian Libraries, University of Oxford.

The most definitive academic study on the subject was conducted by Japanese scholar Dr. Hayashi Takao. Based on factors such as script style, literary form, and mathematical content, he concluded that the manuscript likely dates between the 8th and 12th centuries ([Hayashi, 1995](#), p.23-29). This highlights how even experts can err when interpreting historical documents—but it

is through scientific methods and laboratory analysis that the truth ultimately comes to light. Before the carbon dating of the Bakhshali Manuscript in 2017, the earliest known use of the zero symbol was thought to be the inscriptions of the zero symbol found at Chaturbhuj Temple in Gwalior, Madhya Pradesh, dating to the 9th century CE ([Joseph, 2011](#), p. 254).

In between the Bakhshali Manuscript and the Chaturbhuj Temple's inscription, Aryabhata and Brahmagupta flourished to infuse scientific credence into the mathematical quest of India and the world.

Born in 476 CE in Asmaka on the bank of the Godavari River in Central India, Āryabhaṭa belongs to the same intellectual milieu as the Bakhshālī mathematicians, both representing the remarkable efflorescence of mathematics in India during the early classical period. Āryabhaṭa went to Nalanda University for education and then settled in Kusumapura in modern Patna in Bihar, which was a flourishing intellectual and political center under the Gupta Empire. He composed the pathbreaking *Āryabhaṭīya* in Kusumapura around 499 CE, when he was about 23 years old (Aryabhatiya, 1976, p.xvii-xix). Āryabhaṭīya is one of the earliest and most influential scientific treatises of classical India. Written in concise Sanskrit verse, it served as a foundational text in mathematics and astronomy for centuries.

While the *Bakhshālī Manuscript* shows practical arithmetic with a clear use of zero, Āryabhaṭa's *Āryabhaṭīya* elevates mathematics into a theoretical and astronomical framework. In that sense, both traditions developed side by side — one rooted in applied calculation, the other in systematic science. The text also presents early forms of algorithms—systematic, rule-based procedures for calculation. Āryabhaṭa described methods for extracting square and cube roots, solving indeterminate equations, and working with trigonometric functions through sine tables, all expressed in algorithmic style. These procedures anticipate the algorithmic reasoning that later became central to mathematics and computer science.

Āryabhaṭa introduced a novel method of expressing numbers through an alphabetical code, demonstrating an advanced understanding of positional notation that would later crystallize into the numeral system we now call the Hindu–Arabic numerals. Although later Indian scholars refined the symbol for zero, the structural necessity of zero within positional arithmetic was already inherent in Āryabhaṭa's system (Aryabhatiya, 1976, p. xxvii).

What is truly intriguing is that all major evidence regarding the origin and development of zero exclusively points towards Indian sources. Ancient Indian scholars pioneered the concept of zero—not merely as a placeholder but as a number in its own right—long before it appeared elsewhere. The formal mathematical treatment of zero was first articulated by the Indian mathematician Brahmagupta in the 7th century.

This revolutionary concept then spread from India through the Arab world to Europe, fundamentally shaping modern mathematics and science ([Colebrooke, 1817](#), p.346). The uniquely Indian origin of zero marks a profound intellectual achievement and highlights India's pivotal role in the history of global knowledge.

Table 3. The journey of mathematical numbers.

Text	Approximation of π / or other numbers	Value	Method/Rule	Estimated Period	Source

Text	Approximation of π / or other numbers	Value	Method/Rule	Estimated Period	Source
Rhind Mathematical Papyrus, also known as the Ahmes Papyrus	$\pi \approx 3.1605$	3.1605	Equates the area of a circle to that of a square whose side equals $\frac{8}{9}$ of the diameter	c. 2000 – 1650 BC	British Museum reference numbers Papyrus BM 10057 and BM 10058
Baudhāyana	$\pi \approx 3.088$	3.088	Rule for constructing a circle of equal area to a square: radius = diagonal $\times \frac{13}{15}$	c. 800–500 BCE	Sen and Bag 1983
Āpastamba	$\pi \approx 3.125$	3.125	Improved construction for circle-square area equivalence	c. 600–300 BCE	Sen and Bag 1983
Mānava	$\pi \approx 3.2$	3.2	Simpler fractional approximation	c. 500–300 BCE	Saraswati Amma 1979
Euclid	shows that the areas of circles are proportional to the squares of their diameters. Symbolically, if two circles have diameters d_1 and d_2 , then their areas satisfy:	This relation assumes a constant ratio between the circumference (or area) of a circle and its diameter, which is precisely what we today call π .	Common divisor of two numbers	c. 300 BC	Euclid's Elements of Geometry, by H.M. Taylor, Cambridge University Press, 1893

$$\frac{A_1}{A_2} = \left(\frac{d_1}{d_2} \right)^2$$

Text	Approximation of π / or other numbers	Value	Method/Rule	Estimated Period	Source
Archimedes	Must be between 3.1408 and 3.1429	π = between 3.1408 and 3.1429	Using the method of exhaustion—a forerunner of integral calculus—estimated the ratio of a circle's circumference to its diameter by inscribing and circumscribing regular polygons within and around the circle	c. 287 BC	Archimedes, translated by Thomas Little Heath, Cambridge University Press, 1897
Pingala	Pingala's śūnya → conceptual void → Brahmagupta's 0 (7th c. CE) → numerical zero	Binary	Placeholder for absence in metrical position, Śūnya (शून्य) = “empty” or “void”	c. 2nd century BCE	Chandaḥśāstra
Aryabhata	$\pi \approx 3.1416$	3.1416	The rule is procedural and algorithmic, giving operations rather than abstract constants	c. 5th century CE	<i>Āryabhaṭīya</i> (Gaṇitapāda, verse 10) ¹⁹

$\pi = \frac{\text{Circumference}}{\text{Diameter}} = \frac{62,832}{20,000} = 3.1416$

As explained above with detailed analysis, like other ancient cultures, early Indian mathematics began with number words (*eka* = one, *dvi* = two, etc.) and tally-like reckoning. The Vedas (c. 3500–800 BCE) already attest to a remarkably sophisticated numerical vocabulary, extending to very large magnitudes—*lakṣa* (100,000) and *koṭi* (10 million)—demonstrating that the concept of

scalable numbers was deeply embedded in thought. Yet these were spoken forms, not written numerals.

The use of written numerals emerged centuries later with the Brahmi numerals seen in the 1st century BCE in the Western Deccan Buddhist cave inscriptions. It was the first true Brāhmī numeral signs that introduced graphic symbols for 1 through 9. The use of zero was peculiar $\text{𑀧} = 10$, $\alpha = 20$, $\text{𑀓} = 100$ in Brāhmī numeral signs ([Sircar, 1965](#), p.216).

The systematic series of Brāhmī numerals from 1 to 9 was developed between the 3rd and 1st centuries BCE, but their earliest written representation appears only in the 1st century BCE, in the Western Deccan cave inscriptions²⁰ at sites such as Nasik, Karle, and Kanheri ([Dani, 1963](#), p. 33).

Contrary to popular belief, the Aśokan inscriptions of the 3rd century BCE do not use separate numeral symbols (what we now call Brāhmī numerals, such as $\text{𑀧} = 8$ or $\alpha = 20$). Instead, numbers are written out in words, spelled in Prakrit using the Brāhmī script—for example, *vīsati* (twenty) or *chaḍḍa* (six)²¹. It is possible that written practices of numeral symbols were already emerging during the Aśokan era, or that the quest to devise such symbols gained momentum at that time. However, within the inscriptions themselves, numbers are consistently represented in words rather than through distinct numeral signs. These typically occur in regnal years, counts of animals, or fractions in land/produce records. ([Hultzsch, 1925](#), p.50-174).

The notion of zero, both as a symbol and as a number in its own right, emerged only later in Indian mathematics. Its evolution can be traced through distinct milestones: Pingala's *Chandaḥśāstra* (c. 200 BCE) introduced *śūnya* (“emptiness” or zero) as a combinatorial marker; the Bakhshali Manuscript used a dot (•) to denote zero in calculations; Aryabhata developed the place-value system though without an explicit zero symbol; and Brahmagupta in 7th century CE finally gave zero full numerical and operational status, thereby transforming mathematics marking a decisive breakthrough in the history of mathematics.

Brahmagupta, often hailed as the father of modern mathematics, redefined numerals, ennobled zero as a number, and endowed the decimal system with enduring value. Born in Bhinmal in modern Rajasthan, he wrote his magnum opus ‘*Brāhmasphuṭasiddhānta*’ or ‘The Correctly Established Doctrine of Brahma’ in 628 CE. In this text he catapulted zero from the notion of a placeholder, as in Āryabhaṭīya and the Bakhshālī Manuscript, to treat zero as a number in its own right, subject to defined arithmetic operations. Brahmagupta explicitly described how zero behaves under addition and subtraction, and attempted rules for multiplication and division involving zero.

For example:

- $a+0=a$
- $a-0=a$
- $a \times 0=0$ ²²

Division by zero, he struggled with, describing it but not fully resolving it — a challenge that remained problematic until modern mathematics ([Colebrooke, 1817](#), p.346). He gave systematic methods for solving quadratic, indeterminate, and linear equations, and advanced techniques for handling negative numbers. The *Brāhmasphuṭasiddhānta* not only consolidated Indian mathematics but also codified zero as a mathematical entity, laying the foundation for the full decimal place-value system.

This was the birth of zero as a mathematical entity—a revolutionary concept.

Its transmission through Arabic translations in the 8th–9th centuries played a decisive role in the global spread of the numeral system we use today.

Discovery of zero and development of numeral and decimal systems did not mark an end to mathematical inquiry. Rather, they acted as catalysts that continually augmented and sustained mathematical refinement. It inspired subsequent scholars and thinkers to expand and reimagine the discipline. After Brahmagupta, it was Lalla, who flourished in circa 720 CE, continued the legacy of his predecessors. He was an important but sometimes overlooked figure in the Indian mathematical-astronomical tradition. Nearly nothing is known about Lalla's personal life except some speculative information that he was from Avanti or Malwa in central India. *Śiṣyadhīvrddhidatantra*, or 'A Treatise Which Expands the Intellect of Disciples' and *Jyotiṣaratnakośa* or 'Treasury of Astronomical Lore' were his main works, which systematized Brahmagupta's mathematics ([Pingree, 1981](#), p.89). Lalla's work clarified operations with zero and negative numbers, expanded trigonometry and algebra, and refined astronomical models.

It was the illustrious Śrīdhara of circa 750 CE, who continued the flame of mathematical tradition after Brahmagupta and Lalla. Very little is known about Sridhara's life except his most famous texts, *Pāṭīgaṇita* or *Pāṭīgaṇitasāra*, which deals with practical arithmetic, and *Triśatikā*, also called *Gaṇitasārasaṅgraha*, a concise collection of 300 verses covering a wide range of arithmetic and algebraic topics. He provided systematic rules for operations with fractions, extraction of square and cube roots, computation of interest, and mensuration areas and volumes. Al-Bīrūnī, the great polymath of Khwārizm, travelled to India after the Ghaznavid conquest in circa 1017 CE along with Mahmud of Ghazni. He stayed in India for several years, studying Sanskrit texts and interacting with Indian scholars. His works on arithmetic and astronomy, such as *al-Qānūn al-Mas'ūdī*, cite Indian mathematicians like Āryabhaṭa, Brahmagupta, and Śrīdhara ([Pingree, 1981](#), p. 219).

Mahāvīra was the next great mathematician in the line of great mathematicians who adorned the land of India in the 9th century CE. He belonged to the Jain intellectual tradition and flourished in the 9th century CE during the reign of the Rāṣṭrakūṭa king Amoghavarṣa, who ruled from 814 to 878 CE, in Karnataka. Written around 850 CE, *Gaṇitasārasaṅgraha* is his chief surviving contribution. Unlike Brahmagupta's work, it is devoted purely to mathematics, not astronomy, marking a shift in Indian mathematical literature. He initiated discussion of permutations and combinations, including the formula for combinations, though expressed in verse form, not symbolic notation (Rangacarya, 1912, p.212).

7 INDIAN NUMERALS BECOME ARABIC NUMERALS

The profound mathematical achievements of ancient India—particularly the invention of the decimal place-value system, the concept of zero, and sophisticated techniques in arithmetic, algebra, and trigonometry—were seminal contributions to human knowledge. These innovations did not remain confined to the subcontinent; through channels of trade, scholarship, and translation, they reached the Islamic world. Arab scholars eagerly assimilated these ideas, recognizing their practical and theoretical value. They adopted Indian numerals and integrated Indian methods into their own mathematical works. Far from mere transmission, this process enabled the Arabs to extend, refine, and disseminate Indian mathematical genius across the broader world. Thus, India's original insights became the foundation of the global mathematical tradition, mediated and popularized through Arab scholarship.

The Indian concept of algorithmic steps, as seen in Āryabhata's arithmetic rules, was later codified by Muḥammad ibn Mūsā al-Khwārizmī, a Persian scholar whose work bridged into

Islamic and eventually European scholarship. He was active during the period from circa 780 and 850 CE. Al-Khwārizmī authored *Kitāb al-Mukhtaṣar fī Hisāb al-Jabr wal-Muqābala* or ‘The Compendious Book on Calculation by Completion and Balancing’ in 820 CE in Baghdad, in which he introduced systematic methods for solving linear and quadratic equations. This work laid the foundation for algebra (Khwarizmi, 1342, p.2). The term *algebra*, indeed, itself is derived from *al-jabr*, a word found in the book's title. When his writings were translated into Latin, his name—difficult for Greek and Latin speakers to pronounce—was rendered as *Algoritmi*, from which the term *algorithmus* emerged, eventually evolving into the modern word *algorithm* ([Rosen, 1831](#), p.18).

The next figure in the transmission of the Indian mathematical tradition to Europe was Abū'l-Ḥasan Aḥmad ibn Ibrāhīm al-Uqlīdisī, the distinguished Syrian mathematician of the 10th century. Active in Damascus around 952 CE, he authored the *Kitāb al-Fuṣūl fī al-ḥisāb al-hindī* or ‘*Book of Chapters on Hindu Arithmetic*’, one of the earliest surviving Arabic texts devoted to Hindu arithmetic. His epithet, *al-Uqlīdisī*—meaning ‘the Euclidean’—derives from *Uqlīdis*, the Arabic rendering of Euclid. In this treatise, al-Uqlīdisī systematized Indian numerals and arithmetic operations, adapting them for everyday use in trade, taxation, and record-keeping. His work shows that by his time, the Indian innovations of place-value notation, decimal arithmetic, and the use of zero were no longer foreign curiosities but integral to the Islamic intellectual world. Through scholars like al-Uqlīdisī, the legacy of Indian mathematics was preserved, refined, and made ready for its eventual transmission to Europe, where it would revolutionize calculation and lay the groundwork for modern mathematics ([Saidan, 1966](#), p.3).

Indian numerals spread to the Arab world through trade and scholarship. Arabs referred to these numerals as *hindsa*, meaning ‘Indian’ or *hind-se* ‘from India’ ([Datta and Singh, 1962](#), p.17). This terminology emphasized their origin. From the Islamic world, the numerals reached Europe via Latin translations of Arabic works in Spain and Italy in the 12th century. In particular, Leonardo of Pisa, also known as Fibonacci, in his *Liber Abaci* published in 1202 introduced them widely in Europe. By the 12th century, they became known in Europe as the “Arabic numerals”, though their origin was fully Indian (Lofkar, 2009, p.54).

Fibonacci played a decisive role in the European adoption of the Hindu–Arabic numeral system. Fibonacci, whose birth name was Leonardo Bonacci, was born around 1170 in Pisa, Italy. His father, Guglielmo Bonacci, was a wealthy merchant and customs official stationed at a trading post in Bugia, modern-day Béjaïa, Algeria. As a child, Fibonacci accompanied his father to Bugia, where he received instruction in the Hindu–Arabic numerals and advanced arithmetic from Arabic educators. After his time in Bugia, Fibonacci travelled across the Mediterranean, visiting regions such as Egypt, Syria, Greece, Sicily, and Provence, where he studied the various mathematical systems employed by traders and scholars (Sigler, 2003, p.3-4).

But he never visited India, the land of his mathematical fascination. His mathematical learning originated from the Islamic world, where Indian numerals and mathematical knowledge had already been integrated and were being transmitted through Arabic scholarship—not through direct contact with India.

Fibonacci expressed profound respect for Indian mathematics. He wrote an autobiographical prologue for his seminal *Liber Abaci*, addressing that Michael Scott, his mentor, was a philosopher and scholar with scientific interests in the court of Frederick II. A patron of mathematics and science as well as arts, Frederick II was the Emperor of the Holy Roman Empire in the 13th century. Scott was also mentioned in an uncomplimentary term in Dante's *Inferno*, the first part of *Divine Comedy*.

In this autobiographical prologue, Fibonacci wrote

‘As my father was a public official away from our homeland in the Bugia customhouse established for the Pisan merchants who frequently gathered there, he had me in my youth, brought to him, looking to find for me a useful and comfortable future, there he wanted me to be in the study of mathematics and to be taught for some days (Sigler, 2003, p.15).’

Fibonacci admitted that

‘There from a marvellous instruction in the art of the nine Indian figures, the introduction and knowledge of the art pleased me so much above all else, and I learnt from them, whoever was learned in it, from nearby Egypt, Syria, Greece, Sicily and Provence, and their various methods, to which locations of business I travelled considerably afterwards for much study, and I learnt from the assembled disputations (Sigler, 2003, p.15).’

He admitted not being able to travel to India, but he collected mathematical wisdom from every possible source from the Arab world. He further wrote,

‘But this, on the whole, the algorithm and even the Pythagorean arcs, I still reckoned almost an error compared to the Indian method. Therefore strictly embracing the Indian method, and attentive to the study of it, from mine own sense adding some, and some more still from the subtle Euclidean geometric art, applying the sum that I was able to perceive to this book, I worked to put it together in XV distinct chapters, showing certain proof for almost everything that I put in, so that further, this method perfected above the rest, this science is instructed to the eager and to the Italian people above all others, who up to now are found without a minimum (Sigler, 2003, p.16).’

As we can see, Fibonacci was still using the Roman numerals X and V.

Liber Abaci introduced Indian numerals to the Latin-speaking world and demonstrated their immense practical advantages over Roman numerals and the abacus. By showing how place-value notation and the use of zero simplified arithmetic, bookkeeping, and commercial transactions, Fibonacci provided Europe with a far more efficient mathematical language. His text also extended to algebraic methods and famous numerical examples—such as the sequence later named after him—that revealed the system's broader intellectual potential. Through *Liber Abaci*, the Indian innovation of positional decimal notation, transmitted via Arabic scholars, became firmly established in Europe, laying the mathematical foundation for the Renaissance, the scientific revolution, and ultimately, modern computation.

Leonardo uses the Latinized Arabic word zephirum and not the word nihil for zero. This shows the Arabic source of Leonardo's number zero. We have chosen to translate with the word zephir instead of zero to emphasize this. The modern English word zero is derived from zephir (Boncompagni, 1857, p.5).

The first known European mathematical work in French to use the digit zero appeared in 1275. This was in the *Algorisme* or ‘*Livre de l’algorisme*’, a vernacular adaptation of earlier Latin texts on arithmetic that had been derived from Arabic sources. The appearance of this text in French is significant, for it marks the earliest known use of zero in a European vernacular, rather than in scholarly Latin. In this way, the 1275 *Algorisme* represents a milestone in the diffusion of the

Hindu–Arabic numerals into European culture, bridging the scholarly tradition with the broader intellectual life of medieval France (Smith, 1911, p.128).

The authorship of the French vernacular text *Algorisme* is generally linked to the work of Alexandre de Villedieu, a Norman scholar and poet. Alexandre composed the Latin verse treatise *Carmen de Algorismo* or ‘The Poem on Algorism’, which explained the Hindu–Arabic numerals and the rules of calculation. Around 1275, this text was translated and adapted into French verse as the *Livre de l’algorisme*, making it the first known European vernacular work to employ zero. While the French version is anonymous, its content is directly based on Alexandre's Latin poem, placing him at the root of the *Algorisme* tradition (Leopold, 1894, p.518).

8 CONTINUING QUEST OF INDIAN MATHEMATICIANS

While Arab scholars borrowed extensively from India's mathematical genius and acted as intermediaries in transmitting this knowledge to Europe, the Indian tradition itself did not stagnate. Indian scholars continued to refine, expand, and deepen mathematical inquiry across centuries. Mathematical genius produced original works that advanced algebra, geometry, trigonometry, and astronomical calculations. Their texts introduced novel methods, clarified earlier concepts, and offered systematic approaches to problem-solving. This sustained tradition demonstrates that Indian mathematics was not merely a static heritage exported abroad, but a dynamic, evolving body of knowledge that consistently pushed the frontiers of human intellectual achievement.

After Mahavira, the next prominent mathematician was Āryabhaṭa II. He flourished around 950 CE, possibly in central or northern India. *Mahāsiddhānta* and *Karaṇasāra* were his surviving texts, where he provided sine tables of great accuracy, continuing the Indian development of trigonometric functions ([Dikshit, 1896](#), p.95). Unlike Mahāvīra, who wrote a purely mathematical treatise, Āryabhaṭa II's focus remained tied to astronomy and calendrical science, though mathematical tools formed its backbone.

Jayadeva followed Āryabhaṭa II in mathematical tradition, who was holding the fort during circa 1000 CE. He was a very important mathematician. He is sometimes credited with developing or refining the chakravāla method, a cyclic algorithm for solving indeterminate quadratic equations of the form ([Ramasubramanian, Srinivas, and Sriram, 1994](#), p. 185). This method was later perfected and made famous by Bhāskara II, but scholars acknowledge Jayadeva as an earlier contributor. He inherited the tradition of Śrīdhara and Mahāvīra in arithmetic and algebra.

The next in the line of noted mathematicians was Sripati, who is believed to have lived in the Deccan region of India, possibly in modern-day Karnataka or Maharashtra, during 1019 – 1066 CE. His major mathematical works included *Siddhāntaśekhara*, *Gaṇitatilaka*, *Dhīkotida-karana*, and *Jyotiṣaratnamālā*. He expounded arithmetic, algebra, astronomy and further clarified rules of zero and negative numbers. He improved upon earlier computational techniques ([Datta and Singh, 1962](#), p.90).

The mathematical quest in medieval India continued with the emergence of Hemacandra. He was a Jain polymath from Gujarat, renowned as both a religious leader and a pioneering scholar of mathematics, linguistics, and philosophy. Born in Dhandhuka near Ahmedabad and lived during 1089–1172 CE, he later became associated with the court of King Kumarapala of the Chaulukya dynasty, who deeply respected him for his scholarship and moral influence. In mathematics, Hemacandra is celebrated for his work in combinatorics, particularly his formulation of the

sequence that later came to be known in Europe as the Fibonacci sequence, which he used to describe the number of patterns in poetic meters ([Singh, 1936](#), p.606).

In his *Chandaśāstra* (Treatise on Prosody), written around 1150 CE, he described how to count the number of possible patterns of syllables of short (1 beat) and long (2 beats) length in Sanskrit poetry. This naturally yields the recurrence relation:

$$F(n) = F(n-1) + F(n-2)$$

It was exactly the Fibonacci sequence, centuries before Fibonacci's *Liber Abaci* in 1202 expounded. For this reason, scholars often refer to it as the Hemachandra sequence in Indian mathematics.

His systematic study of permutations and combinations extended earlier Indian traditions and influenced subsequent scholarship. Beyond mathematics, he produced encyclopaedic works on grammar *Siddha-Hema-śabdanuśāsana*, on prosody *Chandaśāstra*, and on Jain philosophy, earning him the title *Kalikalasarvajña* or ‘the all-knowing one of the Kali age’. His personal life reflected the ideals of Jain ascetic discipline, but he simultaneously served as a royal advisor, symbolizing the close ties between intellectual and spiritual life in medieval India.

Bhāskara II or Bhāskarāchārya, who was born in 1114 and lived until 1185 CE, was a native of Bijapur district in present-day Karnataka. His magnum opus *Siddhānta Śiromaṇi* or *Crown of Treatises*, which he published in 1150 CE, is in four parts: *Līlāvātī* or ‘*Arithmetic and Geometry*’, named after his daughter; *Bījagaṇita* or ‘*Algebra*’; *Grahaṇita* or ‘*Mathematics of the planets or astronomy*’, and *Golādhyāya* or ‘*Spherical astronomy*’ revolutionized mathematical tradition in India (Bhāskaracharya, 1816, p.1). He is one of the most celebrated mathematicians in Indian history. Renowned for his extraordinary contributions during the 12th century, he is especially famous for his mathematical treatise *Līlāvātī*, which blends deep mathematical insight with poetic elegance. Although some modern analysts consider him a compiler due to his synthesis of earlier works, Bhāskara II did far more than compile. He refined, extended, and innovated across multiple domains of mathematics. His major contributions are as follows:

Table 4. Bhāskarāchārya's field of work.

Field	Contribution
Number Theory	Chakravāla method for $x^2 - Ny^2 = 1$
Calculus (early ideas)	Rates of change, infinitesimals, turning points
Algebra	Solving quadratic, cubic, and quartic equations
Geometry	Area, volume, cyclic quadrilaterals
Trigonometry	Use of sine, cosine, and angle calculations
Infinity and Zero	Defined concepts of division by zero and infinity
Pedagogy	Mathematical storytelling via <i>Līlāvātī</i>

After Bhāskarāchārya, the next important mathematician was Śāṅkara Vāriyar, who was present during the era 1230 CE. He was a mathematician and astronomer from Kerala, belonging to the rich scholarly tradition that flourished in South India after Bhāskara II. He is best known for his detailed commentaries on Bhāskara's *Līlāvātī* and *Bījagaṇita*, in which he not only explained but

also rationalized the procedures with illustrative examples, thereby strengthening the pedagogical value of these works. His own text, the *Yuktidīpikā* or ‘Lamp of Reasoning’, highlights his commitment to justifying mathematical results with logical proofs rather than relying on rote rules, a distinctive feature that paved the way for the Kerala school's later emphasis on rigorous reasoning in trigonometry and astronomy ([Larson and Bhattacharya, 1987](#), p.2). He was associated with the Nambudiri scholarly tradition and lived in Kerala, where temple-based education and astronomical studies were closely intertwined. Though little is recorded about his personal life beyond his scholarly pursuits, his works reflect a blend of mathematical precision and philosophical reflection, establishing him as a bridge figure between Bhāskara II and the Kerala mathematicians who would eventually develop early forms of calculus.

Bhāskarāchārya inspired generations of mathematicians. The next great mathematical thinker was Gaṅgādhara, a 15th-century mathematician and astronomer. He is best known for his commentary on Bhāskara II's *Līlāvātī*, a foundational text in Indian mathematics. We don't know much about his family, but it is recorded that he was born in a family where generations of scholars endowed the Kerala mathematical tradition. His work, often referred to as *Ganitāmṛta* or ‘The Nectar of Mathematics’, expanded upon Bhāskara's methods in arithmetic, algebra, and geometry, while also illustrating their practical applications in areas such as interest calculation, mixtures, and combinatorics ([Saraswati, 1873](#), p.2). In keeping with the pedagogical style of his time, Gaṅgādhara presented clear explanations and step-by-step reasoning, making his commentary valuable for both teaching and scholarly transmission. Situated historically between Bhāskarāchārya and the Kerala school pioneers, such as Mādhava, who were present during the 14th–15th centuries, Gaṅgādhara represents an important link in the continuity of mathematical learning. *Līlāvātīvivaraṇa* is Gaṅgādhara's commentary on Bhāskara II's *Līlāvātī*, representing one of the key medieval efforts to preserve classical Indian mathematics through explanation, examples, and instruction.

Sūryadāsa was the successor of Gaṅgādhara in mathematical tradition. He was flourishing in the late 15th to early 16th centuries. He is best remembered for his lucid commentary on Bhāskarāchārya's *Līlāvātī*, titled *Līlāvātīvivṛti*. Belonging to the rich tradition of scholastic transmission in medieval India, Sūryadāsa sought to make Bhāskara's compact and often terse verses more accessible to students by providing detailed explanations, worked examples, and practical applications. His commentary illustrates problems of arithmetic, algebra, and mensuration in a pedagogical style that shows concern for teaching and comprehension. In doing so, he ensured the continuity of mathematical learning at a time when classical Sanskrit texts were central to education across India. By situating himself in the lineage of commentators who elaborated upon earlier canonical works, Sūryadāsa played a crucial role in keeping the *Līlāvātī* a living text, bridging the gap between Bhāskara's original genius and its reception among learners several centuries later ([Datta and Singh, 1962](#), p.2).

Kṛṣṇa Daivajña, a prominent 16th-century mathematician and astronomer, is the celebrated author of the astronomical text *Bījagaṇita-vivaraṇa*. His works reveal proficiency in algebra, arithmetic, and astronomical calculations, often framed in ways that could be applied to calendrical and astrological practices. Gaṇeśa Daivajña was the next eminent mathematician and astronomer, who was born in 1507 CE. He was the son of Kṛṣṇa Daivajña. He composed several influential works, the most notable being the *Buddhivīlāsini*, a celebrated commentary on Bhāskara II's *Līlāvātī*. His mathematical interests extended beyond arithmetic and algebra into astronomy, as reflected in his works *Grahalāghava* and *Grahalāghavīya-tīkā*, which provided concise astronomical tables widely used by scholars and practitioners for calendrical and astrological

computations ([Joseph, 2003](#), p.23). Through his writings, Gaṇeśa occupies a significant place in the transmission and elaboration of the Indian mathematical–astronomical tradition, standing at the intersection of scholastic commentary and applied computation. His works ensured the enduring influence of Bhāskara II's legacy well into the early modern period. Rāmakṛṣṇa was another 16th-century Indian mathematician who continued the commentary on Bhāskaracharya's *Līlāvati*.

9 THE DELICATE EMBRACE OF INDIAN MATHEMATICAL TRADITION AND AI

The Egyptian and Indian traditions of practical calculation converged with the Greek insistence on logical rigour, forging a dual legacy: one pragmatic, the other theoretical. This synthesis is precisely the dual character that computer science continues to embody today. Artificial Intelligence draws its essence from multiple ancient streams: (a) algebra and algorithms, inherited through Indian and Arabic transmissions; (b) formal logic, rooted in Greek philosophy; (c) numerical methods, echoing the pragmatic ingenuity of Egyptian and Indian traditions; and (d) geometry and optimization, resonant with the altar-construction rules of the *Śulva Sūtras*. AI, therefore, is not a sudden creation of the modern age, but the continuation of a 5,000-year intellectual lineage—an unbroken evolution of problem-solving strategies, logical structures, and algorithmic reasoning.

In a deeper sense, every time AI ‘learns’ through algorithms, it reenacts the very impulse that once guided a *Śulva Sūtra* scholar measuring ropes to square a circle, an Egyptian scribe calculating the slope of a pyramid, or a Greek philosopher proving a theorem. The continuity lies in humanity's enduring effort to systematize thought, to transform intuition into explicit, repeatable processes—the essence of both mathematics and machine intelligence. Thus, Artificial Intelligence should not be seen merely as a twentieth-century breakthrough, but as the culmination of a profound intellectual journey that began in the ancient worlds of India, Greece, and Egypt, where humanity first discovered how to turn thinking into algorithms.

The historical trajectory traced in this study demonstrates that the pursuit of artificial intelligence cannot be understood as a sudden innovation of the twentieth century. Rather, it is the culmination of a long and layered history of human reasoning, algorithmic imagination, and symbolic abstraction—much of which can be traced to India's mathematical heritage. From the Vedic concern with numbers and geometry, to the sophisticated combinatorial systems of Pingala, and the elevation of *śūnya* to a full numerical status by Brahmagupta, Indian scholars articulated conceptual breakthroughs that resonate directly with the logic of modern computation. The embedding of mathematics into ritual practice, astronomy, and prosody ensured that algorithmic reasoning was not marginal but integral to intellectual life. These traditions preserved, refined, and transmitted an orientation toward structured problem-solving that foreshadowed the rule-based logics at the heart of present-day AI.

Equally significant is the continuity of development across civilizations. Arab mathematicians, such as al-Khwārizmī, systematized Indian arithmetic into al-jabr and, in the process, generated the very word ‘algorithm.’ Their role as custodians and transmitters of Indian genius to Europe ensured that what began in the ritual fire-altars and poetic meters of India reached the Renaissance and, eventually, modern science. Figures like Fibonacci explicitly acknowledged the superiority of the Indian method over Roman numerals and abacus-based calculation, marking a decisive shift in

European thought. The genealogy of AI, therefore, must be situated not in the laboratories of mid-twentieth-century America alone but within this millennia-long history of cross-civilizational mathematical imagination.

10 CONCLUSION

Indian scholarship itself did not stagnate once its innovations were transmitted abroad. On the contrary, the tradition flourished continuously. From Mahāvīra's combinatorics to Bhāskaracharya's proto-calculus and Hemacandra's poetic permutations that anticipated the Fibonacci sequence, Indian thinkers consistently expanded the conceptual repertoire of mathematics. Their efforts reveal that discovery was never treated as an endpoint but as a catalyst for sustained inquiry. This recursive pattern—innovation, refinement, transmission, and renewal—mirrors the very logic of artificial intelligence, which thrives not on isolated breakthroughs but on iterative training, feedback, and improvement. AI is better understood, therefore, as the latest expression of a deep civilizational tendency toward abstraction, systematization, and algorithmic reasoning.

The global story of AI is, then, one of layered inheritances. The symbolic void of *śūnya* became the binary foundation of Pingala's metrical system, which in turn resonates with the zero-one logic of modern computing. The systematic place-value notation, perfected in India and mediated through the Islamic world, became the indispensable architecture for calculation without which no computer could function. The recursive methods of Indian mathematicians prefigured the very logic of modern programming, in which problems are decomposed into smaller steps, rules, and conditional operations. What appears today as an advanced field of machine learning is, in essence, a continuation of humanity's age-old quest to mechanize thought, encode procedure, and formalize reasoning.

To regard artificial intelligence merely as a twentieth-century invention is thus to truncate history and obscure the intellectual debts owed to earlier civilizations. The Indian mathematical imagination, with its synthesis of symbolic representation, algorithmic reasoning, and cosmological speculation, provided fertile ground for the emergence of computational thinking long before machines existed. AI is not an abrupt departure but a sustained unfolding of ideas whose roots lie deep in human cultural and intellectual traditions.

By reclaiming this long *durée* perspective, we are reminded that technological modernity is inseparable from ancient knowledge systems. The trajectory from Vedic ritual geometry to digital neural networks underscores the continuity of intellectual effort across epochs and geographies. Artificial intelligence is, in this light, not simply a machine-age phenomenon but the latest chapter in humanity's enduring engagement with numbers, forms, and algorithms. It is the outcome of a civilizational dialogue across millennia—an inheritance enriched by Indian mathematics, shaped by Islamic scholarship, and carried forward in European science—now converging in a global pursuit to replicate intelligence itself.

In this sense, AI should be understood as the most recent crystallization of a sustained human project: the quest to formalize knowledge, render thought calculable, and extend the power of reasoning beyond the human mind. Far from being an isolated achievement of modernity, it is a testimony to the cumulative genius of civilizations and to the resilience of ideas first articulated thousands of years ago on the Indian subcontinent.

NOTES

1. सहस्रशीर्षा पुरुषः सहस्राक्षः सहस्रपात् । स भूमिं विश्वतो वृत्वा अत्यतिष्ठदशाङ्गुलम् ॥ Rgveda, Maṇḍala 10, Sūkta 90. [↗](#)
2. कासीत्प्रमा प्रंतिमा किं निदानुमाज्यं किमासीत्परिधिः क आसीत् । छन्दः किमासीत्प्रउंगं किमुक्थं यद्देवा देवमयजन्तु विश्वे ॥ Rgveda, Mandala 10, Sukta 113, Sloka 3. [↗](#)
3. द्वादश प्रथयश्चक्रं एकं त्रिणि नाभ्यानि क उतच्चिकेत । तस्मिन्शतमिह परि षक्त आसीत्सश्चक्र आहवना अजगन्नाम ॥ Rgveda, Mandala 1, Sukta 164, Sloka 22. [↗](#)
4. *Rig Veda*, tr. by Ralph T.H. Griffith, 1896, Mandala 1, Sukta 90, Hymn 3 and 4. [↗](#)
5. Krishna Yajurveda, Mandala 3, Sukta 7, Hymn 7. [↗](#)
6. एकोऽसि त्रयः पञ्च सप्त नवैकादश त्रयोदश पञ्चदश सप्तदश नवदशैकविंशतित्रयोविंशतिः पञ्चविंशतिः सप्तविंशत्यः नवविंशत्यःैकत्रिंशच्च त्रयस्त्रिंशच्च ॥ Shukla Yajurveda, Adhyaya 17, Sloka 2. चत्वारि चाष्टौ च द्वादश च षोडश च विंशतिश्च चतुर्विंशच्च अष्टाविंशच्च द्वात्रिंशच्च षट्त्रिंशच्च चत्वारिंशच्च चतुश्चत्वारिंशच्च अष्टचत्वारिंशच्च ॥ Shukla Yajurveda, Adhyaya 17, Sloka 3. [↗](#)
7. एकं दश शतम् सहस्रम् अयुतम् नियुतम् प्रयुतम् अर्बुदम् न्यर्बुदम् समुद्रः मध्यः अन्तः परार्धः इत्य् एतानि नामानि भवन्ति ॥ Yajurveda, Vājasaneyi Saṃhitā, Chapter 17, Sloka 2. [↗](#)
8. In Śānti Parva (Book 12, Chapter 231) of the *Mahābhārata*, where Bhīṣma describes the structure of the Yuga cycle and the cosmic calculations of time. Hymn 19 - I shall, in their order, tell you the number of years that are for different purposes calculated differently, in the Krita, the Treta, the Dwapara, and the Kali yugas. Hymn 20 - Four thousand celestial years is the duration of the first or Krita age. The morning of that cycle consists of four hundred years and its evening is of four hundred years. Hymn 21 - Regarding the other cycles, the duration of each gradually decreases by a quarter in respect of both the principal period with the minor portion and the conjoining portion itself. Hymn 29 - The learned say that these twelve thousand celestial years form what is called a cycle. A thousand such cycles form a single day of Brahman [(Brahmā)]. Hymn 30 - The same is the duration of Brahman's [(Brahmā's)] night. With the beginning of Brahman's day the universe begins to come into being. During the period of universal dissolution the Creator sleeps in Yoga meditation. When the period of sleep expires, He awakes. Hymn 31 - What is Brahman's day covers a thousand such cycles. His night also covers a thousand similar cycles. They who know this are said to know the day and the night. [↗](#)
9. In the *Mahābhārata*, the Pāṇḍavas were five brothers, while their cousins, the Kauravas, numbered a hundred. Both groups were locked in a fierce struggle for the throne of the Kuru dynasty; परस्परविरोधे तु वयं पञ्चशते शतम् । परैस्तु विग्रहे प्राप्ते वयं पञ्चाधिकं शतम् ॥ *Mahābhārata*, Sabhāparvan (Book of the Assembly Hall), Adhyāya 26, Verse 7. [↗](#)
10. “śikṣā kalpo vyākaraṇaṃ niruktaṃ chando jyotiṣaṃ iti vedāṅgāni”, (*Nirukta* 1.20). [↗](#)
11. ‘Deva-yajñāya śulbaṃ vyākhyāsyāmaḥ’, We shall now explain the rules of the *śulva* (measuring cord) for the sacrifice to the gods. [↗](#)
12. “Alpakṣaram asandigdham sāravad viśvatomukham | astobham anavadyam ca sūtram sūtravido viduḥ ||” “A *sūtra* is defined by those who know: concise in words, unambiguous, full of essence, universal in scope, without padding, and faultless.” [↗](#)
13. दीर्घचतुरस्रस्याक्षया रज्जुः पार्श्वमानी तिर्यग्मानी च यत् पृथग्भूते कुरुत् तदुभयं करोति ॥ [↗](#)
14. समस्य द्विकर्णि प्रमाणं तृतीयेन वर्धयेत् । तच्च चतुर्थेनात्मचतुस्त्रिंशोनेन सविशेषः ॥ [↗](#)
15. Heath notes that Pythagoras's contributions were “concerned with the properties of whole numbers and their ratios ... and there is no evidence of any treatment of the measurement of circles or the ratio of their circumference to diameter.” [↗](#)

16. Papyrus BM 10057 and BM 10058, British Museum, United Kingdom. [↗](#)
17. Book XII, Proposition 2. [↗](#)
18. Brahmasphuṭasiddhānta, Chapter 18, Sloka 30–34. [↗](#)
19. चतुरधिकं शतमष्टगुणं द्वाषष्टिस्तथा सहस्राणाम्। अयुतद्वयस्य विभक्तं ऋणं शेषं दशाधिकं शतम्॥ [↗](#)
20. – = 1; = = 2; ≡ = 3; ¥ = 4; h = 5; € = 6; ? = 7; ∩ = 8; † = 9; α = 10; Θ = 20; ∞ = 30; 8 = 40; G = 50; J = 60; L = 70; O = 80; ⊕ = 90; H = 100; T = 200; . = 300; (up to 900 using additive forms); \ = 1000; 2 = 2000; (and so on, additive/multiplicative patterns used). [↗](#)
21. देवानं पियस पियदसिन लाजिन विसविसाभिषितेन अट्टविसतिभिसेहि अट्टविसति॥ [↗](#)
22. शून्येन गुणिता राशयः शून्य एव भवेत्। शून्येन द्वौ योगौ शून्य एव। शून्यात् शून्यं व्यावर्त्यते शून्यमेव। शून्येन विभक्तं शून्यमेव। शून्येन विभक्तः संख्या शून्यहरसमाप्ता। शून्यं शून्येन विभक्तं शून्यमेव। Brahmasphuṭasiddhānta, Chapter 18, Sloka 30–34. [↗](#)

REFERENCES

- Abraham Seidenberg (1978). *The Origin of Mathematics*. Baltimore: Johns Hopkins University Press.
- Archimedes, translated by Thomas Little Heath, Cambridge University Press, 1897
- Aristotle. (1928). *The Works of Aristotle Translated into English: Volume I – Organon I* (Categories, On Interpretation, Prior Analytics). Trans. Harold P. Cooke and Hugh Tredennick. Oxford: Clarendon Press. [↗](#)
- Arthur Anthony Macdonell, *A Vedic Grammar for Students: Including a chapter on syntax and three appendices—list of verbs, metre, accent*, Clarendon Press (Oxford), 1916. [↗](#)
- Arthur Berriedale Keith, *The Religion and Philosophy of the Veda and Upanishads*, Harvard University Press, Cambridge, Massachusetts, 1925. [↗](#)
- Aryabhatta, *Aryabhatiya*, edited by Kripa Shankar Shukla, New Delhi: Indian National Science Academy, 1976.
- Aryabhatta, *Aryabhatiya: An Ancient Indian Work on Mathematics and Astronomy*, translated by Walter Eugene Clark, Chicago: University of Chicago Press, 1930.
- B.K. Thapar. (1975). *Kalibangan: A Harappan Metropolis Beyond the Indus Valley Expedition*, Vol. 17, No. 2. [↗](#)
- Bakhshali Manuscript, MS. Laud Misc. 671, Bodleian Libraries, University of Oxford.
- Banerji, Sures Chandra. *A Companion to Sanskrit Literature: Spanning Four Thousand Years*. 2nd edn revised and enlarged edition. Delhi: Motilal Banarsidass, 1963.
- Baudhāyana Śulva-sūtra. Edited with commentary by Pt. Sudhakara Dvivedi. Benares: Medical Hall Press, 1890.
- Bühler, G. (1896). *Indian Palaeography*. Strassburg: Karl J. Trübner.
- Burkert, Walter. (1972). *Lore and Science in Ancient Pythagoreanism*. Cambridge, MA: Harvard University Press.
- Chandah Sutra of Pingala Acharya, edited by Pandita Visvanatha Sastri, Calcutta: 1874, Asiatic Society of Bengal publication.
- Colebrooke, H.T. (1817). *Algebra, with Arithmetic and Mensuration, from the Sanscrit of Brahmagupta and Bhāscara*. London: John Murray. [↗](#)
- Confirms Usage of Numeral Symbols in Western Cave Inscriptions Dated to the Late 1st c. BCE.*
- D. Srinivasachar and Vidwan S. Narasimhachar Edited, *Apastamba-Sulvasutra*, Mysore: Government Branch Press, 1931. [↗](#)

- D.E. Smith, History of Mathematics, Vol. II; Ettore Picutti, Sul Numero e la sua Storia.
- Dani, A.H. (1963). *Indian Palaeography*. Oxford University Press. [↗](#)
- Dani, S.G., (2003). 'On the Pythagorean triples in the Śulvasūtras'. *Current Science*, 85(2). [↗](#)
- Datta, B. and Singh, A.N. (1935), Title: *History of Hindu Mathematics*, Part I (The Śulba Sūtras)
Publisher: University of Lucknow.
- Datta, B., and Singh, A.N. (1935). *History of Hindu Mathematics: A Source Book (Part II: Algebra, Trigonometry, and Calculus)*. Lahore: Motilal Banarsidass. [↗](#)
- Datta, Bibhutibhusan, and Singh, Avadesh Narayan, (1962), *History of Hindu Mathematics, Part I: Numerical Notation and Arithmetic*. Asia Publishing House. [↗](#)
- Dikshit, S.B. (1896). *Bhāratiya Jyotisha*. Poona. English translation by R. V. Vaidya. 2 pts. New Delhi: Government of India Press, Controller of Publications, 1969, 1981. [↗](#)
- Dvivedin, Sudhkara (ed. and comm.) (1910). *Mahāsiddhānta*. Benares Sanskrit Series Vol. 36, nos. 148–150. Benares. (Reprint, New Delhi: Caukamba Sanskrit Prathista, 1995.)
- Euclid's Elements of Geometry, by H.M. Taylor, Cambridge University Press, 1893
- F.R. Allchin and B. Allchin. (1982). *The Rise of Civilization in India and Pakistan*. Cambridge: Cambridge University Press, 1982. [↗](#)
- Fibonacci's Liber Abaci: A Translation into Modern English of Leonardo Pisano's Book of Calculation*, translated by Laurence Sigler, New York: Springer, 2003
- Folkerts, The Early History of the Indo-Arabic Numeral System: The Arabic Sources, 1996, and Georges Ifrah, The Universal History of Numbers, 1998.
- Gillings, R.J. (1972). *Mathematics in the Time of the Pharaohs*. MIT Press, Cambridge, MA. [↗](#)
- Hannah Devlin, Much Ado About Nothing: Ancient Indian text contains earliest zero symbol, Guardian, 14 September 2017. [↗](#)
- Hayashi, T. (1995). *The Bakhshali Manuscript: An Ancient Indian Mathematical Treatise*. Groningen: Egbert Forsten. [↗](#)
- Hayashi, T. (2003). *Indian Mathematics*. In: Flood, G. (Ed.), *The Blackwell Companion to Hinduism*, Blackwell. [↗](#)
- Heath, Thomas L. (1921). *A History of Greek Mathematics, Vol. I: From Thales to Euclid*. Oxford: Clarendon Press. [↗](#)
- Hultzsch, E. (1925). *Inscriptions of Asoka. Corpus Inscriptionum Indicarum*, Vol. I. Oxford: Clarendon Press. [↗](#)
- Joseph, G.G. (2011). *The Crest of the Peacock: Non-European Roots of Mathematics* (3rd ed.). Princeton University Press. [↗](#)
- Joseph, G.G. (2003). "Different ways of knowing: Contrasting styles of argument in indian and greek mathematical traditions." In *Mathematics Education and Philosophy*, Routledge [↗](#)
- K.V. Abhyankar and V.P. Limaye (eds.), (1938). *Upaniṣad-saṅgraha*, Poona: Ānandāśrama Sanskrit Series, No. 97.
- Lakshman Sarup (ed. and trans.), (1920). *The Nighaṇṭus and the Nirukta of Yāska*, Harvard Oriental Series, Vol. 5, Cambridge. [↗](#)
- Lal, B.B. (2002). *The Earliest Civilization of South Asia: Rise, Maturity and Decline*. New Delhi: Aryan Books International. [↗](#)
- Larson, G.J., and Bhattacharya, R.S. (1987). *Sāṃkhya: A Dualist Tradition in Indian Philosophy*. Princeton University Press. [↗](#)
- Lavis, A. (2024). *The Origin and Significance of Zero: An Interdisciplinary History*. Routledge. [↗](#)
- Leonardo of Pisa Fibonacci, *Liber Abaci*, ed. Baldassarre Boncompagni, 1857, Rome: Tipografia delle Scienze Matematiche e Fisiche, Vol. 1.

- Leopold Delisle, Manuscripts Latins Et Francais, Paris: H. Champion, Libraire, 1894, 518
- Lofker, K. (2009). *Mathematics in India*. Princeton University Press.
- M. Tr. Rangacharya, *The Ganita-sara-sangraha of Mahaviracarya*, Madras: Government Press, 1912.
- Mahābhāṣya of Patañjali on *Pāṇini's Aṣṭādhyāyī* (Kielhorn ed., 1880), Vol. I
- Morris Kline (1972). *Mathematical Thought from Ancient to Modern Times*. Oxford University Press. [↗](#)
- Muhammadibn Musa al-Khwarizmi, *Kitab fi al-Jabr wa al-Muqabala*, 1342.
- Olivelle, Patrick (ed. and trans.). *Dharmasūtras: The Law Codes of Āpastamba, Gautama, Baudhāyana, and Vasiṣṭha*. Annotated Text and Translation. Delhi: Motilal Banarsidass, 1999.
- Pāṇini, *Aṣṭādhyāyī with Kāśikā*, Vol. II, ed. S.D. Joshi and J.A.F. Roodbergen, Poona: Deccan College, 1991. [↗](#)
- Pāṇinīya Siksā Or The Siksā Vedanga, Manomohan Ghosh, Calcutta: University Of Calcutta, 1938
- Pingree, D. (1981). *Jyotiḥśāstra: Astral and Mathematical Literature*. Wiesbaden: Harrassowitz. [↗](#)
- Proclus Lycaeus (tr. Glenn R. Morrow) (1970). *A Commentary on the First Book of Euclid's Elements*. Princeton University Press.
- R.J. Gillings (1971). “The rhind mathematical papyrus: Its place in the history of mathematics.” *The Mathematics Teacher* 64(7). [↗](#)
- Ramasubramanian, K., Srinivas, M.D., and Sriram, M.S. (1994). Mathematics of the planet earth: The cyclic chakravala method. *Indian Journal of History of Science*, 29(2). [↗](#)
- Rath, Saroj Kumar, (2025). *India-Afghanistan: From Common Inheritance to Contested Histories*, London: Routledge Publication. [↗](#)
- Rosen, F. (Trans.). (1831). *The Algebra of Mohammed Ben Musa*. London: Oriental Translation Fund. (First English translation of al-Khwārizmī's algebra). [↗](#)
- Saidan, A.S. (1966). *The Arithmetic of Al-Uqlīdisī: The Story of Hindu-Arabic Arithmetic as Told in “Kitāb al-Fuṣūl fī al-ḥisāb al-hindī”*. Dordrecht: D. Reidel. [↗](#)
- Sarasvati, S.S. (Ed.) (1982). *Mānava Śulba Sūtra*. Delhi: Indian Council of Historical Research. [↗](#)
- Saraswati, S. (Ed.). (1873). *Līlāvātī with the Gaṇitāmṛtasāgara ṭīkā of Gaṇḍādhara*. [Rare Book Archive, Benares Edition]. [↗](#)
- Seidenberg (1962). “The ritual origin of geometry.” *Archive for History of Exact Sciences* 1(5). [↗](#)
- Sen, S.N., and Bag, A.K. (1983). *The Śulbasūtras and the Arithmetical Combinatorics of Prosody*. *Indian Journal of History of Science*, 18(1).
- Sen, S.N. and Bag, A.K., 1983. *The Śulba Sūtras of Baudhāyana, Āpastamba, Kātyāyana and Mānava*. New Delhi: Indian National Science Association. [↗](#)
- Singh, A.N. (1936). On the Use of Fibonacci Numbers in Indian Mathematics. *Osiris*, 1. [↗](#)
- Sircar, D.C. (1965). *Indian Epigraphy*. Motilal Banarsidass. [↗](#)
- Smith, D.E., and Karpinski, L.C. (1911). *The Hindu-Arabic Numerals*. Boston: Ginn & Co.
- T. Eric Peet (1923). *The Rhind Mathematical Papyrus: British Museum 10057 and 10058*. Clarendon Press, Oxford.
- Talageri, Shrikant G. (2000), *The Rigveda: A Historical Analysis*, Aditya Prakashan, New Delhi.
- The sacred laws of the Aryas as taught in the schools of Apastamba, Gautama, Vasishta and Baudhayana, Translated by George Buhler, Part-II Vasishta and Baudhayana, Oxford: The Clarendon Press, 1882.
- Thibaut, George, ed. and trans., *The Baudhāyana Śulvasūtra*, Bibliotheca Indica, 1875. [↗](#)
- Thomas L. Heath (1956 [orig. 1908a]). *The Thirteen Books of Euclid's Elements* (Vol. 1). Cambridge Publications.

- Thomas L. Heath (Ed. and Trans.) (1956 [orig. 1908b]). *The Thirteen Books of Euclid's Elements* (Vol. 3). Cambridge Publications.
- W. Caland translation of Baudhayana Sruta Sutra, Vol.1, Calcutta: Baptist Mission Press, 1902.
- Will Durant, *The Story of Civilization, Vol. I: Our Oriental Heritage*, New York: Simon and Schuster, 1935. [📄](#)
- Winternitz, Moriz. *A History of Indian Literature. Vol. I: Introduction, Veda, Epics, Purāṇas*. Translated from German by V. S. Sarma. Reprint of the 1927 edition. Delhi: Motilal Banarsidass, 1981.
- Yash Pal, Baldev Sahai, R.K. Sood, and D.P. Agrawal, *Remote Sensing of the 'Lost' Sarasvati River*” Proceedings of the Indian Academy of Sciences (Earth and Planetary Sciences) 1980.

A Critique of ChatGPT: The Challenge of Large Language Models to the Enlightenment Project

Germán Bula

University of Bogota, Columbia

DOI: [10.1201/9781042018116-5](https://doi.org/10.1201/9781042018116-5)

ABSTRACT: This paper explores the potential threat that Large Language Models (LLMs) pose to the Enlightenment project, as defined by Immanuel Kant: the cultivation of autonomous thought. While outsourcing trivial cognitive tasks to LLMs may seem harmless, this practice risks forming habits that undermine intellectual autonomy. Two key concepts guide this inquiry: *bullshit*, in Harry Frankfurt's technical sense—discourse indifferent to truth—and *laziness*, as Kant understands it, particularly in its political form: the unwillingness to assume responsibility for one's thoughts. LLMs, by design, generate plausible but non-committal text—bullshit, which can foster a climate increasingly tolerant of such discourse. Moreover, the habitual outsourcing of thinking can transform mere laziness into what can be called political laziness, dulling the impulse to form and defend one's views. Unlike tools with limited functions (e.g., calculators), LLMs are general-purpose cognitive agents, and this universality risks expanding their influence across domains where critical judgment is vital. As LLMs are increasingly used for sensitive or existential decision-making (e.g., therapy), a positive feedback loop emerges in which intellectual responsibility and care for truth become ever more eroded. This paper argues that such trends are not merely technological shifts but represent a serious departure from the Enlightenment ideal.

Keywords: Kant, enlightenment, LLMs, laziness, bullshit, and habit

1 METHODOLOGY

This paper adopts a conceptual–critical methodology grounded in philosophical analysis and normative critique. It undertakes close textual engagement with Kant's writings on Enlightenment and laziness, alongside Frankfurt's analytic account of *bullshit*, and brings these frameworks into dialogue with the functional design and use-patterns of contemporary Large Language Models. Rather than empirical testing, the study proceeds through conceptual clarification, comparative reasoning, and reflective analysis of illustrative cases, to assess the political and epistemic implications of LLMs for intellectual autonomy and Enlightenment ideals.

2 INTRODUCTION

“The Age of Enlightenment”: these words bring to mind images of well-dressed savants in cafés dreaming up a new world free from tyranny and superstition; and also of men of science working away in makeshift observatories and laboratories, laying the groundwork for the breathtaking advances in science and technology we see today. In the age of Rousseau and Newton, scientific progress went hand in hand with the project of emancipation through the use of reason (see eg., [Freudenthal, 1986](#)). However, this association between the Enlightenment and technology may be historical and accidental; indeed, the Enlightenment is a political project, and one can envision highly technological societies that are far removed from the Enlightenment ideas (e.g, the current fears about “techno-fascism”, [Chayka, 2025](#)). Not only can technology thrive in unenlightened societies, it is also possible that certain technologies are, in their nature, hostile to the Enlightenment project. This paper explores this possibility in relation to Artificial Intelligence (AI) in general, and Large Language Models (LLMs) in particular.

Worries about a clash between the Enlightenment project and AI are not new. In 2018, after the AI program Alpha Go defeated the world's greatest Go players with hitherto unknown strategies, but before the 2022 launch of ChatGPT, former U.S defense secretary Henry [Kissinger \(2018\)](#) wrote an essay for the *Atlantic* entitled *How the Enlightenment Ends*. Two of Kissinger's worries are relevant to this essay: 1) the opacity of self-training AI with regards to human understanding: that is, the fact that the solutions produced by AI may not be rationally reconstructed by the human intellect, and 2) the capacity of AI to modify human patterns of thought (as exemplified by human players adopting Alpha Go's novel strategies)

Kissinger's first worry echoes that of Husserl's Vienna lecture (1996, 49-74) that as scientific knowledge accumulates, it can become a mass of rote (though effective) procedures divorced from intellectual understanding (a seeing, in consciousness, why such and such a proposition is necessary). For both Kissinger and Husserl, it is not enough to reap the technical or informational benefits of science, but it is important for humans to understand their origins. I will argue i) that this imperative of understanding is central to the enlightenment project and ii) is directly harmed by a tendency of LLM use to promote a tolerance for *bullshit*, in the precise technical sense defined by [Frankfurt \(2005\)](#): a use of discourse that has no regard for its truth or falsity (and, therefore, for its coherence, rationality or authenticity).

Kissinger's second worry will be echoed when we come to the topic of laziness. I will argue, based on what we know of the incorporation of tools into bodily schemas (e.g [Clark, 2004](#)), habitual use of LLMs s can lead to a habit of off-loading cognitive tasks that includes not only rote cognitive tasks (which would be harmless to the Enlightenment project) but also those of a higher order, such as the forming of political opinions.

I will further argue that these two tendencies are synergistic, and can lead to a positive feedback loop in which intellectual responsibility and care for truth become progressively more eroded. As will be discussed later, there is an important internal relationship between bullshit and laziness: bullshit (be it in academic, interpersonal or political contexts) eschews the work of authenticity, and can therefore be born of laziness. But, more crucially, this form of labour-saving can become habitual, so that what was initially, say,

inauthentic bullshit for the sake of efficiency or political expediency can become laziness, as a habitual disposition.

My arguments on the impact of LLMs on the enlightenment project will be formulated against the backdrop of Kant's *What is Enlightenment?*, which I take to be a representative formulation of the Enlightenment project; and which I will discuss in the first section. In Section 2, I will clarify my view on the nature of LLMs. This will lead to a discussion of bullshit in Section 3: what it is, how LLM generated discourse corresponds to the definition, how it has become pervasive in our world, how a climate of tolerance for bullshit is self-reinforcing, and how LLM discourse contributes to this trend. In Section 4 I will discuss LLMs as a labour-saving device, and what kind of uses can be sensible from the point of view of an economy of effort. However, I will complicate this discussion by exploring the consequences of habitual use of LLMs, and how habit can lead us to use LLMs in ways that betray the enlightenment project. Section 5 concludes the chapter by discussing the synergy between bullshit and laziness, and pondering how to respond to the threat that LLMs pose to the Enlightenment project.

3 WHAT IS ENLIGHTENMENT?

In 1784, the *Berlinische Monatschrift*, in response to controversies regarding secularization, launched a call for essays on the question “What is Enlightenment?”. Kant's answer, both pithy and philosophically weighty, has become classic:

Enlightenment is the human being's emergence from his self incurred Minority. Minority is inability to make use of one's own understanding without direction from another. This minority is self incurred when its cause lies not in lack of understanding but in lack of resolution and courage to use it without direction from another. Sapere aude! Have courage to make use of your own understanding! is thus the motto of enlightenment. ([Kant, 1996](#), p.11).

Enlightenment is a matter of courage and resolution; bearing in mind the religious controversies that form the backdrop of the essay, Enlightenment can be called a properly *spiritual* matter. Whatever social conditions facilitate or hinder Enlightenment, there is a key role for the individual, a *resolution* to formulate (and live by) one's own opinions according to reason.

Enlightenment, as understood by Kant, goes beyond freedom of speech and a puerile “not being told what to think”; making use of one's own *understanding* means taking care of thinking rational (as opposed to capricious) thoughts. This is well illustrated in Kant's *Dreams of a Spirit Seer*, a philosophical takedown of the mysticism of Emmanuel Swedenborg. Swedenborgian mystics enjoy a kind of freedom in that, untethered by empirical considerations, their individual fabrications can run wild: “The land of shadows is a paradise of dreamers. Here they find unlimited country where they may build their houses *ad libitum*” ([Kant, 1900](#), p.37). This, however, does not mean that they are enlightened. A further effort is needed; to purify oneself from fancies and wishful thinking, and to weight reasons for and against:

I have purified my soul from prejudices, I have destroyed any blind affection which ever crept in to procure in me an entrance for much fancied knowledge (...) nothing is venerable to me but what enters by the path of sincerity into a quiet mind open to all reasons be thereby my former judgment confirmed or abolished, be I convinced or left in doubt (...) The judgment of him who refutes my reasons, is my judgment, after I first have weighed it against the scale of self-love, and, afterwards, in that scale against my presumed reasons, and have found it to have a higher intrinsic value. ([Kant, 1900](#), 85).

In *What is Enlightenment*, Kant identifies to causes whereby we fail to do this kind of work ([Kant, 1996](#), p.11): laziness and cowardice. As illustrated by Kant's discussion of Swedenborg, autonomous thought requires effort. Why does it require courage as well? There are at least two reasons: 1) thinking for oneself makes one *responsible* for one's positions; this responsibility can no longer be offloaded to an external guardian “It is so

comfortable to be a minor!” ([Kant, 1996](#), p.11; see [Altemeyer 2006](#), 122-124 for a description of the hedonic benefits, especially reduced anxiety, of offloading decisions to an authorized guardian such as a priest), and 2) once one has decided that a certain belief or course of action is rational, one can only see divergence from it as a kind of self-betrayal (see [Kant, 1997](#), pp.40-41); that is, one is compelled by self-respect to follow through with one's rational beliefs, which may be dangerous or inconvenient (e.g, opposing slavery in a society that practices it).

In a trivial sense, LLMs do some thinking for us (e.g, suggest travel plans, write business letters, etc). But performing such rote tasks does not seem to entail a danger to the enlightenment idea of thinking for oneself in a higher, spiritual sense (e.g, deciding on one's religion or political beliefs). In what follows, I will argue that both laziness and cowardice, as well as a climate of tolerance for bullshit, are all mutually self-reinforcing, in such a way that even seemingly innocent uses of LLMs can help to jeopardize the Enlightenment project. But first, it is important to understand what LLMs *are*.

4 WHAT IS AN LLM?

A tempting framework for looking at LLMs is that of *dual use*: technologies can be used for good and bad purposes (see e.g, [Vaynman and Volpe, 2023](#), for the case of military vs civilian uses), therefore it is the use of technologies, rather than their nature, that should be the focus of analysis. However, LLMs have certain peculiarities that complicate such a framework: 1) LLMs are not designed, like most tools, to facilitate human labour, but rather to replace it (Kirschlaeger, 2025), even when the replacement service or product is of markedly inferior quality. 2) LLMs are importantly misunderstood; there is much hype surrounding their current and near-future technical possibilities (see [Thais, 2024](#); for the emergence of magical thinking around new technologies, see Thurschwell, 2001).

For Peter Kirschlaeger (2025), the term *AI* is a misnomer when applied to LLMs; and the term *data-based systems* would be more appropriate. Indeed, LLMs do not solve problems when asked questions (e.g, they do

not have an internal model of the problem situation and try out different solutions to see which works best); rather, they make use of their training data to come up with statistically likely answers (this explains why they have trouble with mathematical problems, [Vinyes, 2024](#)).

Understanding what LLMs do can help in delineating the type of problems in which LLMs perform well from those in which they perform poorly. Since LLMs do not solve problems like humans do, the set of problems where they perform well or ill does not neatly map on to problems that humans find easy or difficult. As an example, the Monty Hall is frustratingly difficult to grasp for many humans: in a game-show setting, a player is given the choice of opening one of three doors for a prize: behind two are goats, behind the other is a car. Once the player has selected a door (say, door 1), the host opens another (door 2) and reveals there was a goat behind it, and then asks the player whether he would like to change his initial choice. Should she? It is hard for human minds to understand that the probabilities for door 1 are 33%, while those for door 3 are 50%: we tend to think the odds for either remaining door are the same. ChatGTP 3.5 did not suffer from such cognitive dissonance, and confidently gave the right answer. However, it gave the same “right” answer when the initial conditions of the problem were altered (e.g, the player is a goat enthusiast who has no use for a car), since these unusual twists were not present in the training data ([Morse, 2023](#)). This example illustrates the precise manner in which it can be said that ChatGTP does not know what it is talking about (and not knowing what one is talking about is a good definition of bullshit, as we shall see in the next section).

There is, however, another twist to the story. After ChatGPT's troubles with variations on the Monty Hall problem were publicized, these were included in the training data of new versions. In the sense of giving appropriate answers, no problem is difficult for LLMs, provided the answers are included in their training data; but there is no hard and fast rule to know what is or is not included. Therefore, the “expertise” of LLMs has a jagged frontier (cfr. Dell’acqua *et al.*, 2023). LLMs perform certain tasks well, and certain tasks poorly (but always give confident sounding answers, even for their hallucinations, see [Gómez-Rodríguez, 2025](#), 74), and sometimes it is not easy to know which tasks are which.

If we accept the premise that LLMs sometimes perform poorly, and that the occasions on which they do so are not easily predictable, we can prove that LLMs, taken in themselves, are not intelligent. The good performance of LLMs depends on a human user who can select tasks for which LLMs are competent, and who can detect when the LLM has made a mistake: therefore, it is the system composed of user-plus-LLM that is intelligent. Another way of characterizing LLMs is that they produce bullshit, in the technical sense of producing the appearance of knowledge rather than knowledge itself. I will explore this idea in the following section, as well as the ways in which bullshit can be detrimental to the Enlightenment project.

5 BULLSHIT

As Harry [Frankfurt \(2005\)](#) explains in a now classic essay, bullshit is not the same as lies: a liar has truth in mind when she deviates from it, whereas a bullshitter is unconcerned with truth or falsehood. A good example is a student who answers a teacher's question with a rambling word salad: she wishes to give the impression that she knows something she does not (and would be delighted if what she was saying turned out to be true). Another example would be a politician justifying a hard-to-justify decision: rather than trying to push a false story on the public, his aim is to obfuscate, to make it through the interview. In this sense, bullshit is a worse enemy of truth than mere falsehood, in that, in as much as it becomes pervasive, it erodes the habit of verification, of attending to the way in which things actually are ([Frankfurt, 2005](#)) of seeing, within one's consciousness, *how* something is true, as [Husserl \(1996\)](#) calls us to do. In many ways, Frankfurt's discussion of bullshit echoes [Plato's \(1987\)](#) discussion of rhetoric and sophistry in the *Gorgias*: whereas philosophers seek truth and justice, the likes of Gorgias or Pollus strive for the *appearance* of these things; and, crucially, this striving for appearances is not only separate from, but positively detrimental to the striving for the real thing. Indeed, once the impression that a task that needs doing is done, it is harder to actually do it (a good example of this is *greenwashing* by ecologically irresponsible companies, see [Kennedy et al., 2022](#)).

Do LLMs produce bullshit? The results of prompts, of course, may be true or false, useful or harmful, etc, (however, it bears noting that, since they depend on training data, the more there is bullshit in the data, the more likely it is that it will show up on the results). But there is a sense in which LLM results are always bullshit: the programs i) do not know what they are talking about, ii) are designed to give the impression that they do, and iii) have no filter for false or incoherent answers. For example, when confronted with a trigonometry problem, LLMs do not perform calculations on angles or hypotenuses to solve it; rather, they seek likely responses to the prompt among their training data, and present that as an answer. Is the answer correct? LLMs have no way of checking; the only partial guarantee of correctness is the fact that the answer is likely based on the training data. That is, they give the impression of knowledge without concerning themselves with knowledge: this fits Frankfurt's definition of bullshit quite precisely.

Why is this relevant? Besides isolated instances of bullshit, we can also speak of contexts in which bullshit is prevalent. Writing in 1986, [Frankfurt \(2005\)](#) notes that there seems to be a lot of bullshit around, and attributes this to a public sphere that calls upon persons (both public and private) to instantly opine on whatever topic is being discussed at a given time, whether they have knowledge of it or not. This atmospheric pressure towards bullshit is even higher today, when social media demands and rewards instantaneous (and often anonymous, and thus more irresponsible) “hot takes”. But this pressure has not existed at all times. Jhon Henry Newman, in the 1850's, observed with astonishment what he saw as a nascent phenomenon:

An intellectual man, as the world now conceives of him, is one who is full of “views” on all subjects of philosophy, on all matters of the day. It is almost thought a disgrace not to have a view at a moments notice on any question (...) every day there must be a supply, for the gratification of the public, of new and luminous theories (...) the journalist lies under the stern obligation of extemporizing his lucid views, leading ideas, and nutshell truths for the breakfast table. ([Newman, 2008](#), 12-13)

Newman views this exigence and habit of “extempore philosophy” as a “rod of cruel slavery” (2008, 13). This instantaneous opining (that is second-nature to us in the present) caused, for Newman, “a distress sometimes so keen and so specific that it resembles nothing else than bodily pain” (2008, 14). Quite possibly, Newman was especially attentive to truth and authenticity, but the fact that he could write as he did about the evils of instantaneous opining is proof enough that he wrote for an audience that was less used to bullshit than we are in the present.

There is, then, such a thing as *atmospheric bullshit*: contexts where bullshit is more prevalent, and therefore, where expectations of truth and authenticity are lowered. This is important because bullshit works better when it finds a collaborative audience, one that will not examine it closely. Bullshitting one's way through an essay question can be a successful strategy provided that the teacher is too tired, lazy, or indifferent to detect bullshit; a bullshit justification for a political decision needs a complacent media that won't call it out, etc. In climates of high atmospheric bullshit, people will tend to expect bullshit from the messages they receive, and therefore, both to heighten their tolerance for bullshit and to expect high tolerance for bullshit they themselves produce. This heightened bullshit production elevates atmospheric bullshit, in a positive feedback loop.

This is where the prevalence of LLMs comes into play; in a climate where a good number of messages are produced through LLMs, there is a heightened expectation that messages will be produced by a “clanker” ([Jones, 2025](#)) rather than a human being: that is, there will be a lowered expectation of authenticity, responsiveness, and a care for truth. Arwa Mahdawi, who begins an article on the topic by assuring us that she is, indeed, a human being, states:

The internet is rapidly being overtaken by AI slop. (It's not clear who coined the phrase but “slop” is the advanced iteration of internet spam: low-quality text, videos and images generated by AI.) A recent analysis estimated that more than half of longer English-language posts on LinkedIn are AI-generated. Meanwhile, many news sites have covertly been experimenting with AI-generated content – bylined, in some cases, by AI-generated authors.” (Mahdawi, 2025)

Not only are messages produced in an inauthentic manner, but they are also read in this way. In Calvino's *If on a Winter's Night a Traveller* (1982), written in 1979, Calvino presents us with the wealthy fiction writer Silas Flannery, who produces books that mimic successful styles through mechanical algorithmic means; and, a few pages later, with the literary critic Lotaria, who uses computers to speed up her work: rather than reading the books, she consults a computer generated table of word frequencies in a texts, which gives her an idea of what the themes of the book are, and enough to go on for her to produce a critique. What seems to worry Calvino is a coming world in which readers pretend to read and writers pretend to write; that is, a world where neither senders nor receivers of messages attend to the content of the message. Perhaps we are in such a world: not only are students using AI to produce papers, but teachers are using it to grade them ([Murphy, 2024](#)). This economy of unattended-to messages competes with that of authentic messages: [Tullis \(2025\)](#) argues that the proliferation of “AI slop” (that is, low-quality AI-produced content) discourages both writers (who have to stand out over against an enormous mass of cheaply produced content) and readers (who have to sift through this mass of content in search for something of quality).

One may even speak of instances of *metabullshit*: that is, of content that, rather than attempting to appear authentic for a reader (as in first-order bullshit), counts on a reader that will not examine the content at all, and thus the content barely gestures at authenticity, as if saying “here is where authentic content would go, if one were bothered to produce it”. An example of *metabullshit* could be the 1994 interview of musician Brian Ferry published by *Ray Gun* magazine (see [Miranda, 2020](#)): the entire text of the interview was rendered in a font called Zapf Dingbats, wherein all letters are replaced by symbols such as stars, triangles or spades. Of course, the interview is completely illegible: is this not an implicit confession that no one cares about the contents of the interview, but rather about the aesthetic feel of the magazine, the “vibe”? The fact that “AI slop” is recognized as such, and that there is so much of it, indicates that there is a market for filler content, *metabullshit*.

If LLMs help increase atmospheric bullshit, they are inimical to the Enlightenment project. As we saw in our discussion of Kant's *Dreams of a Spirit-seer*, thinking for oneself requires *work*, a care for whether one has

purified oneself of prejudices, and thoroughly weighed arguments for and against a proposition; and bullshit eschews precisely this kind of work. LLMs are a labour-saving device; but the Enlightenment project is only accomplished if we do the work.

6 LAZINESS

In Douglas Adams's *Dirk Gently's Holistic Detective Agency* ([Adams, 1988](#)), written in 1987, we meet an electric monk: a technological advance designed to save users the time and effort of having spiritual beliefs. The particular electric monk that appears in the narrative is damaged, and firmly believes the world is pink. The electric monk is a good symbol for LLMs both in that people use them to save themselves the work of constructing beliefs, and in that they can have amusing hallucinations.

Why would people cede such a core task to an external machine? As a part of a progression:

The Electric Monk was a labour-saving device, like a dishwasher or a video recorder. Dishwashers washed tedious dishes for you, thus saving you the bother of washing them yourself, video recorders watched tedious television for you, thus saving you the bother of looking at it yourself; Electric Monks believed things for you, thus saving you what was becoming an increasingly onerous task, that of believing all the things the world expected you to believe ([Adams, 1988](#), p.3)

As technology advances, we outsource more and more of our tasks to it. Modern people are not only used to giving this or that task to a machine, but have come to expect that a machine will come up to perform whatever task they find tedious. This is not necessarily a bad thing: one finds very little to object to using a calculator to perform a tedious sum.

To be sure, when one does a sum in a calculator, strictly speaking, one is not “thinking for oneself”, but in a way that is not incompatible with the Enlightenment project. The autonomous thinking this project requires has to

do with taking a stance on the issues the conduct of one's life. Indeed, the guardians that Kant decries are books that tell you what to think, spiritual advisors that tell you what to believe and doctors that decide on your regimen ([Kant, 1996](#), p.11). Whether I perform a sum on a calculator or in my head does not affect the autonomous conduct of my life, in which lies my dignity ([Kant, 1996](#), p.22).

What about using ChatGPT to write an essay for some tedious class, or a message of condolences for an acquaintance? Is it not the same sort of situation, where we are feeling lazy and don't wish to spend time on a given task? An important difference is that in an essay, at least putatively, we are taking an intellectual stance on a given subject; and in a message of condolences, we are conveying a feeling and an attitude towards the grieving party that, putatively, we are meant to actually have.

This difference becomes clearer if we distinguish between kinds of laziness. Using a calculator to avoid the task of figuring out how much to tip at a restaurant can be called *energetic laziness*; that is, an unwillingness to spend effort towards a given task. But humans also display *political laziness*, that is, an unwillingness to develop and commit to certain stances, feelings or attitudes; it is a laziness with regards to the autonomous and rational conduct of one's own life (once again, see [Altemeyer, 2006](#), 122-124 for the hedonic payoff of offloading these kinds of decisions to a guardian). The kind of minority that worries Kant has to do with *political* rather than *energetic* laziness; and is closely related to the cowardice that Kant also names as a cause for it: in this sense, laziness and cowardice are not two separate causes of minority, but rather, taken together, a name for the kind of attitude that leads people to offload to others the task of figuring out how to live.

What kind of laziness motivates the use of LLMs? Of course, many uses have to do with energetic laziness; but there are growing signs of political laziness as well, such as the worrying trend of using LLMs for psychological advice (cfr. Moore *et al.*, 2005). Furthermore, as is hinted at in Douglas Adams' progression from dishwashers to electric monks, energetic and political laziness can be causally related: I wish to argue that the habit of using LLMs for trivial tasks can lead to a dependence on LLMs for more important matters.

Let us look at the task of writing an essay. It can be divided into two sub-tasks: i) the substantial task of taking a stance on a subject, and ii) the formal task of writing the essay, which is a matter of presentation: putting together the text and bibliography according to a set of norms (APA or MLA), polishing the writing style, etc. It would appear that it is safe, from an Enlightenment point of view, to give the formal tasks to an LLM, so long as the substantial task is performed by the human in question.

However, the tasks of writing and thinking can't be so neatly divided: thinking through a problem involves writing out ideas in order to see whether they make sense; that it is, we need to reify ideas in order to have them in front of us and view them critically ([Bula, 2008](#)); when a human goes through several drafts of an essay she is usually revising her thinking on the subject, not only her presentation of it.

Habitual offloading of the task of writing unto LLMs can, by and by, turn into a habit of offloading the task of thinking and taking stances. If an LLM is habitually used as an aid to writing, it will become an expected and relied-upon part of the writer's resources (in a similar way in which physical tools become a part of the user's bodily scheme, see [Baccarini et al., 2024](#); [Bahmad, 2020](#); [Clark, 2004](#)); in phenomenological terms, what we do habitually, becomes sedimented as passive synthesis, as part of the ground out of which we constitute experience ([Venebra, 2018](#)). In the process of writing, whenever the user encounters a task that she perceives as rote or tedious, she will unthinkingly open the ChatGPT tab. There is a risk that more and more kinds of tasks will come to seem tedious and off-loadable to ChatGPT, in a self-reinforcing positive feedback loop. ChatGPT may create a slippery slope, in which more and more cognitive tasks are renounced (see [Kosmyna et al., 2025](#), where it is suggested that use of LLMs reduces neural connectivity, memory recall and critical thinking skills). In this way, energetic laziness can become political laziness: we may go from asking LLMs to polish the expression of our thoughts to asking them to think for us.

Habitual use of a calculator carries no such risk. The key difference is that calculators are tools for a specific set of tasks, whereas LLMs can be asked anything (and will come up with some kind of response, adequate or not). Therefore, the habit of asking questions of LLMs can be expanded to all kinds of topics, from specific information to political views and

psychological advice. Calculators can add and subtract for us, but LLMs can lead us to a habit of not thinking for ourselves.

7 LLMS AND THE ENLIGHTENMENT PROJECT

I have endeavoured to show that broad and habitual use of LLMs is inimical to the Enlightenment call for autonomous thought in at least two aspects: i) LLMs increase atmospheric bullshit, and therefore produce a climate in which inauthenticity thrives, and ii) habitual use of LLMs leads to an increase in the offloading of cognitive tasks, both trivial and important. In both aspects, there is a self-reinforcing feedback loop: i) the more bullshit there is in a social system, the more bullshit there will tend to be; ii) the more I offload thinking onto LLMs, the more I will tend to do so in the future, and for more and more tasks. Furthermore, both aspects reinforce each other as well: in a climate of high atmospheric bullshit, people will be more tempted to save time and let LLMs do the thinking for them because there is a lowered expectation for authentic thought, and a lowered chance of finding adequate recipients for it.

I have not discussed other dangers to the Enlightenment project that fall outside the scope of this paper, such as the unthinking tribalism that can be caused by social media (Sánchez, 2023), the growing monopolization and political use of mainstream news sources ([Merrick, 2025](#)), or the return of fascistic political tendencies ([Peters, 2019](#)). If my arguments are correct, habitual and quotidian use of LLMs *contributes* to the imperilment of the Enlightenment project.

If one believes the Enlightenment project to be worth preserving, how should one respond?

Recent publications reflect a kind of fatalism, with many variations on the theme of “AI is here to stay” (e.g., [Skipworth, 2025](#); [Jazbec et al., 2025](#); [Gee 2025](#)), followed by suggestions on how to adapt to it in different fields. While workers in different fields where the Enlightenment project is or should be promoted (e.g., education, science journalism, etc) certainly need to adapt to a changing jobs market and work environment, I would posit that a robust defence of the Enlightenment project as such is needed as well.

Educators can take concrete steps to resist the erosion of autonomous thinking (such as banning the use of electronic devices in the classroom, or insisting on personal, authentic, elements in their assignments), and communicators can fight back against “AI slop” by developing compelling writing styles that are difficult to emulate by LLMs; but a clear articulation of the Enlightenment project and how it is threatened in the twenty-first century can give clarity and focus to these efforts.

If we imagine the Enlightenment as an eighteenth century affair, a matter of men in powdered wigs, we risk seeing it as something already accomplished, and we risk taking it for granted. Not only has humanity never reached an enlightened age where autonomous thinking is the norm, but there is no guarantee that the gains that have been accomplished are stable and permanent. Work must be done to preserve the Enlightenment ideal; this means taking on the Enlightenment as a political project in the present, and defending autonomous thought and the dignity that comes with the rational conduct of one's own life. Hopefully, I have shown that the defense of these ideas is not something antiquated, but an urgent task for our present.

ACKNOWLEDGEMENT

This chapter is a partial result of the research project “El problema de la formación: actualización de un concepto fundamental en la educación y su despliegue en la enseñanza de la filosofía.”, with funds from the Government of Colombia.

REFERENCES

- Adams, D. (1988). *Dirk Gently's Holistic Detective Agency*. New York: Pan Books ↵
- Altemeyer, B. (2006). *The Authoritarians*. Self published: <http://members.shaw.ca/jeanaltemeyer/drbob/TheAuthoritarians.pdf> ↵

- Baccarini, M., Martel, M., Cardenalli, L., Sillan, O., Farnè, A., y Roy, A. (2024). "Tool use imagery triggers tool incorporation in the body schema". *Front Psychol.* 2014 May 30;5:492. <https://doi.org/10.3389/fpsyg.2014.00492>. PMID: 24910624; PMCID: PMC4038856.↵
- Bahmad, S. (2020). "Insights into body representation plasticity following tool-use: the role of vision and proprioception.". *Cognitive Sciences*. Université Lumière Lyon 2, 2020.↵
- Bula, G. (2008). "Algunas ideas lonerganianas para la enseñanza de la escritura académica". *Logos* no.13, pp. 73–80↵
- Calvino, I. (2021). *If on a Winter's Night a Traveler*. New York: Harcourt Brace.
- Clark, A. (2004). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. New York: OUP↵
- Chayka, K. (2025). "Techno-Fascism comes to America". The New Yorker, february 26: <https://www.newyorker.com/culture/infinite-scroll/techno-fascism-comes-to-america-elon-musk>↵
- Dell'Acqua, F., McFowland, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Kraye, L., Candelon, F., y Lakhani, K. "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality". Harvard Business School, Working Paper 24-013, https://www.hbs.edu/ris/Publication%20Files/24-013_d9b45b68-9e74-42d6-a1c6-c72fb70c7282.pdf
- Frankfurt, H. (2005). *On Bullshit*. New York: PUP↵
- Freudenthal, G. (1986). "The Relationship Between Natural and Social Philosophy in the Work of Newton, Rousseau, and Smith". In: *Atom and Individual in the Age of Newton. Boston Studies in the Philosophy of Science*, vol 88. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-4500-5_8↵
- Gee, D. (2025). "AI is here to stay – but do we know it?". The Pie, feb 13, 2025: <https://thepienews.com/ai-is-here-to-stay-but-do-we-know-it/>↵
- Gómez-Rodríguez, C. (2025). "Grandes modelos de lenguaje:¿de la predicción de palabras a la comprensión?" en Betanzos, A., Peña, D., y Poncela P. (eds) *La Inteligencia Artificial hoy y sus aplicaciones con Big Data*. Madrid: FUNCAS↵

- Husserl, E. (1996). *La Filosofía en la crisis de la Humanidad Europea*. Valencia: Universitat de València[↵]
- Jazbec, S., Leva B., and Licardo, M. (2025). “AI Is Here to Stay: An Empirical Study of Attitudes Among Teachers of English and German”. *English Language Overseas Perspectives and Enquiries*. Vol. 22, no. 1: <https://journals.uni-lj.si/elope/article/view/21842>[↵]
- Jones, CT. (2025). “How “Clanker” became the Internet's new favorite Slur” Rolling Stone, august 6, 2025: <https://www.rollingstone.com/culture/culture-features/clanker-cogsucker-robot-ai-slurs-viral-1235401262/>[↵]
- Kant, E. (1900). *Dreams of a Spirit-Seer*. London: Macmillan.[↵]
- Kant, E. (1996). *Practical Philosophy*. New York: CUP[↵]
- Kant, E. (1997). *Groundwork of the Metaphysics of Morals*. New York: CUP[↵]
- Kennedy, Merlita; Dela, Pooja; Sibeko, Ziyanda; Lötter, Bernadette; Ishmael, Jared; Serongoane, Tobia. (2022). The Green, The Bad and The Ugly: The Risks of Greenwashing. *Journal of the Southern African Institute of Mining and Metallurgy*. 122. vi–vii[↵]
- Kirchschlaeger, P. (2025). *Ethics and the Digital Transformation of Human Work*. London: Palgrave Macmillan
- Kissinger, H. (2018). “How the Enlightenment Ends”. The Atlantic, June 2018, <https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/>[↵]
- Kosmyrna, N., Hauptman, E., Ye Tong, Ye, Situ, J., Xian-Hao, L., Beresnitzky, A., Braunstein, I., y Maes, P. (2025). “Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task”. Arxiv.org, <https://arxiv.org/abs/2506.08872>[↵]
- Mahdawi, A. (2025). “AI-generated ‘slop’ is slowly killing the internet, so why is nobody trying to stop it?”. *The Guardian*, Wed 8 Jan 2025 20.50 GMT. <https://www.theguardian.com/global/commentisfree/2025/jan/08/ai-generated-slop-slowly-killing-internet-nobody-trying-to-stop-it>
- Merrick, A. (2025). “How Media Consolidation Affects the News You See”. Chicago Booth Review, march 04, 2025: <https://www.chicagobooth.edu/review/how-media-consolidation-affects-news-you-see>[↵]

- Miranda, P. (2020). “*David Carson: Surfing the Unconventional Waves of Zapfs Dingbats*”. Medium, June 27, 2020: <https://medium.com/@yopaulmiranda/david-carson-surfing-the-unconventional-waves-of-zapfs-dingbats-52a520186c64>
- Moore, J., Grabb, D., Agnew, W., Klyman, K., Chancellor, S., Ong, D., y Haber, N. (2025). “Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers”. The 2025 ACM Conference on Fairness, Accountability, and Transparency en arXiv:2504.18412
- Morse, A. (2023). “LLMs and Reasoning, Part I: The Monty Hall Problem”. Medium, Jun 6 2023. <https://ahmorse.medium.com/llms-and-reasoning-part-i-the-monty-hall-problem-f30b22c7ade7>
- Murphy, S. (2024). “Teachers are Using AI to Grade Essays. But Some Experts are Raising Ethical Concerns”. CNN, Abril 6 2024, <https://edition.cnn.com/2024/04/06/tech/teachers-grading-ai>
- Newman, J.H., (2008). *The Idea of a University Defined and Illustrated*. Project Gutenberg: Ebook 24526. <https://www.gutenberg.org/files/24526/24526-pdf.pdf>
- Peters, M. (2019). “The Return of Fascism: Youth, Violence and Nationalism”. *Educational Philosophy and Theory*, vol, 51, issue 7, pp 674–678
- Plato (1987). *Gorgias*. Indianapolis: Hackett
- Sanchez, P. (2023). “21st century: The Return of Tribalism?” *Instituto Español de Estudios Estratégicos*, jan 11, 2023: https://www.ieee.es/Galerias/fichero/docs_analisis/2023/DIEEEA01_2023_PEDSAN_Tribalismo_ENG.pdf
- Skipworth, C. (2025). “AI Is Here To Stay: How To Adapt, Thrive And Lead”. Forbes, jul 14 2025: <https://www.forbes.com/councils/forbesbusinesscouncil/2025/07/14/ai-is-here-to-stay-how-to-adapt-thrive-and-lead/>
- Thais, S. (2024). “Misrepresented technological solutions in imagined futures: The origins and dangers of AI hype in the research community”. *Proceedings of the Seventh AAI/ACM Conference on AI, Ethics, and Society* (AIES2024). <https://ojs.aaai.org/index.php/AIES/article/view/31737/33904>

- Thurschwell, P. (2021). *Literature, Technology and Magical Thinking 1880-1920*. Nueva York: Cambridge University Press
- Tullis, J. (2025). “*Sifting Through the Slop: How Generative AI Created a Market for Lemons for Text-Based Works*”. Available at SSRN: <https://ssrn.com/abstract=5266660> or <http://dx.doi.org/10.2139/ssrn.5266660>
- Vaynman, Jane; Volpe, Tristan. (2023). “Dual use deception: how technology shapes cooperation in international relations.” *International Organization*. 77. 599–632. [10.1017/S0020818323000140](https://doi.org/10.1017/S0020818323000140).
- Venebra, M. (2018). “Husserl: Cuerpo Propio y Alienación”. *Investigaciones Fenomenológicas* no.15, pp. 109–132
- Vinyes, M. (2024). “Large Language Models and Math: A Review of Approaches and Progress”. *Medium*, Nov 12, 2024. <https://medium.com/@vinyes.marina/large-language-models-and-math-a-review-of-approaches-and-progress-b58c76e7716e>

Artificial Intelligence in Financial Markets: Decoding the Investors' Sentiments

Leena Chhabra

University of Delhi

Niti Nandini Chatnani

Indian Institute of Foreign Trade, New Delhi

DOI: [10.1201/9781042018116-6](https://doi.org/10.1201/9781042018116-6)

ABSTRACT: The fast-evolving world of artificial intelligence has reasonably advanced the study of financial markets and investor psychology with the application of machine learning algorithms in investor sentiment analysis. Individual investors are often driven by cognitive biases while making choices, and their sentiments shape the trajectory of the market. This paper maps the developments in the study of sentiment analysis for comprehending investor behavior and explores the application of machine learning approaches. A systematic inquiry is made by performing a keyword search in the Scopus database to study publication trends, knowledge structure, and extensively review the models used for sentiment analysis. It is found that the machine learning based models are increasingly adopted for pattern identification and sentiment prediction. This inquiry highlights the role of social media in predicting the market outcomes. Findings of this study are useful for regulators to decode investor sentiments for market stability and future researchers to work on voice and video data for sentiment classification.

Keywords: Individual Investor, Investment, Machine Learning, Artificial Intelligence, Financial Market

1 INTRODUCTION

Individual investors in stock markets tend to be naïve and largely trade based on their sentiments ([Odean, 1998](#); [Han and Kumar, 2013](#)). Out of a number of stocks, they are likely to buy or chase a stock that recently grabbed their attention ([Goetzmann and Kumar, 2008](#)). Unlike institutional investors, who are professionals and rely on specialised analytical platforms, individual investors are more likely to make investment decisions influenced by heuristics, emotions, and social discussions ([Haritha P H, 2024](#); [Snow and Rasso, 2025](#)). Such decisions can cause markets to deviate from fundamentals and increase volatility ([Barber and Odean, 2008](#)). Investor sentiments are those beliefs about the future flows of a stock and associated risks that are not substantiated by the present information, and such beliefs or predictions influence investors' decisions to invest in a stock ([Baker and Wurgler, 2007](#)).

Over time, the study of investor sentiments has garnered significant attention due to their impact on investment decisions and market outcomes. Studies examining the sentiment-driven trading decisions of individuals find that the stocks purchased by individual investors perform worse than the stocks they sell ([Odean, 1999](#)). It has been observed that they have a tendency to place high number of limit orders than high net worth individuals and institutional investors (Linnainmaa, 2003). Owing to irrational choices, they often engage in excessive trading and lose money ([Shleifer and Summers, 1990](#)). Studies examining the outcome of such trades find that they increase volatility in the market ([Baig et al., 2022](#)) and due to their contrarian trading behavior, individual investors provide liquidity to the markets against institutional investors by buying when they sell. This function contributes to the stability of markets. Retail trades positively contribute to stock liquidity initially, but this impact reverses in subsequent weeks ([Cheng et al., 2021](#)), and such behavior of individual investors can encourage professional traders to trade against them.

Investor sentiments, though, define the trajectory of the stock market; their quantification and measurement have been challenging for researchers and practitioners. With the onset of artificial intelligence (AI), the study of investor sentiments has changed drastically from qualitative judgements to sophisticated and statistically robust machine learning (ML) based approaches. With the advances in technology, investor sentiments can be studied by the posts, tweets, comments or any opinion shared on social media platforms. Sentiment analysis or opinion mining analyzes the direction of the text. It assesses if the content is objective or subjective and if the subjective content represents positive or negative sentiments ([Schumaker et al., 2012](#)). The sentiments are generally classified as positive or negative. Other categorizations like hopeful, happy, satisfied, optimistic, pessimistic etc. are also possible.

This paper maps the methodological advancements in the study of investor sentiment analysis from early behavioral observations and qualitative assessments to

the so far developed AI applications. It performs a critical review of ML based approaches and their accuracy in sentiment measurement. Additionally, the study goes further to assess the potential of these approaches in the prediction of market outcomes. An integrated approach combining bibliometric mapping and systematic review is followed to comprehensively analyze the domain knowledge and offer policy recommendations. Through an extensive inquiry spanning more than 55 years, the paper offers a critical review of methodologies and data sources.

As distinct from previous studies, this paper is broader in perspective and specific in its approach to focus on methodological developments. In particular, the following unique contributions are made by this study:

- i. A broader period from 1969 to 2025 is considered for this study;
- ii. Publication trends and knowledge structure are identified based on the analysis of bibliographic information using VOSviewer software; and
- iii. A detailed content analysis of select papers on ML approaches is performed to offer a comprehensive understanding of advances in the domain.

2 METHODOLOGY

In line with the objectives of this study, two search strings focusing on (i) sentiment analysis in financial markets, and (ii) sentiment analysis using AI tools in financial markets were employed. The search was conducted in the Scopus database, which is a database for academic literature covering more than 25000 journals, widely accepted in the disciplines of social sciences. The keywords in the first search string included sentiment, financial market, share market and stock market, and the other search string included additional keywords of ML and AI.

The search results were filtered based on language; papers only in the English language were considered due to language barrier and analysis compliant structure. The number of studies considered was 2806 for sentiment analysis in financial markets and 200 for sentiment analysis using AI tools in financial markets. Their respective bibliographic information was exported and processed through VOS viewer software. A comparative analysis is made on the publication trend, disciplinary fragmentation and knowledge structure. Further, the studies based on AI tools are described for their publication source and country collaboration to show the global research network.

An extensive review of systematically selected papers is performed to study the methodological approaches used in the study of investor sentiments. The selection

criterion of these papers is based on PRISMA guidelines([Page et al., 2021](#)), as explained in the following flow chart:

[Figure 1](#) explains the paper selection process adopted for the review. Initially, a keyword search was conducted in the Scopus database to retrieve relevant results. These papers were further shortlisted based on language and type of study. The refined search produced 179 results. Of which, 101 papers were selected based on Title review, and the remaining 78 papers were found irrelevant to the theme and rejected. In the next step, an abstract review was performed. Out of these 101 papers, 62 were identified as suitable, and they were thoroughly reviewed.

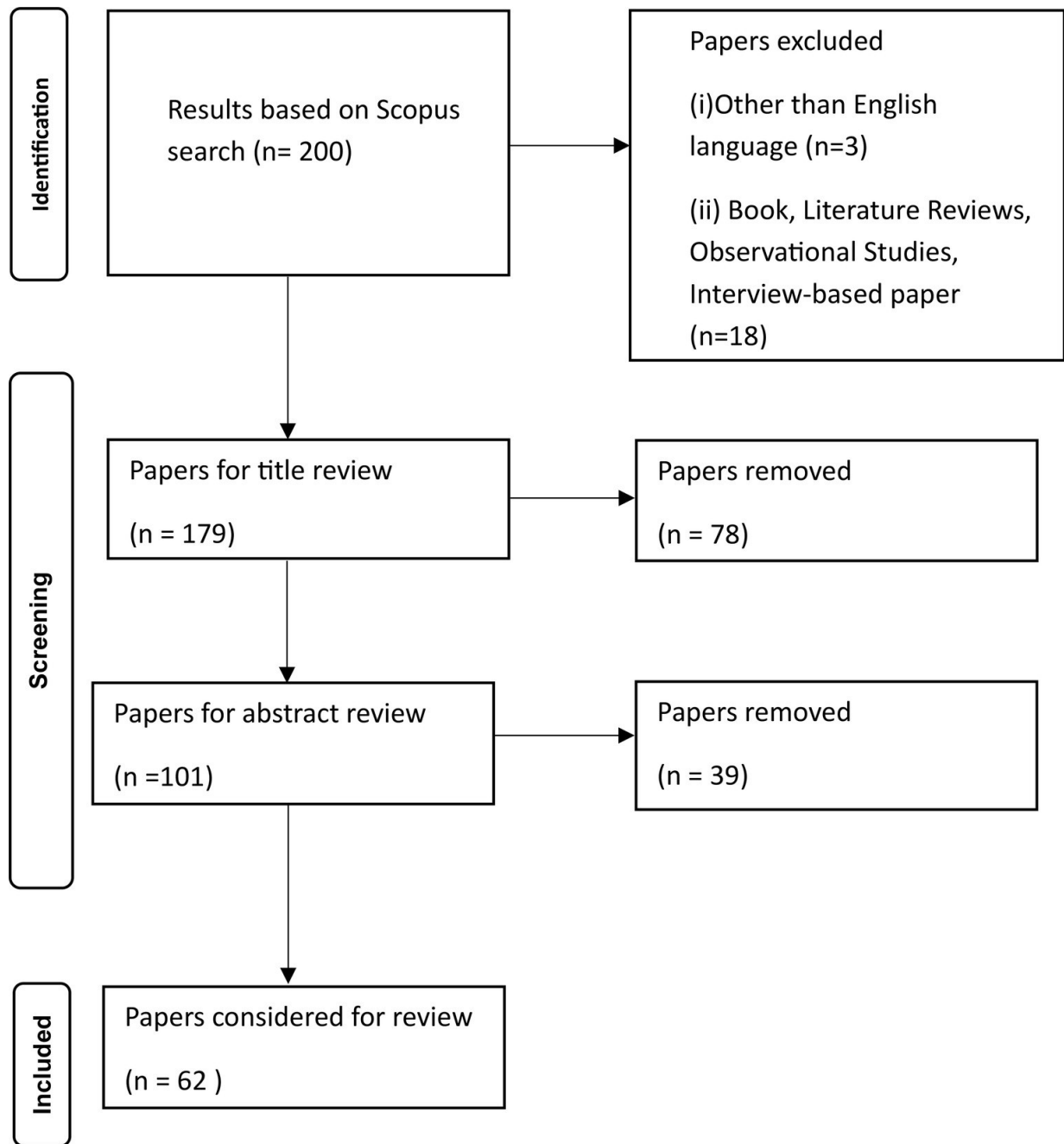


Figure 1. Selection process for Review based on PRISMA guidelines. [📄](#)

The exclusion criteria

Literature reviews, Interview-based papers, observational studies, books and conference papers were excluded for review.

3 RESULTS

3.1 Publications trend

The study of sentiments in financial markets has grown substantially since its first publication in 1996. [Figure 2](#) highlights the rising number of publications on sentiment analysis within the stock market context, but the application of AI in sentiment analysis has been observed since 2009 ([Figure 3](#)).

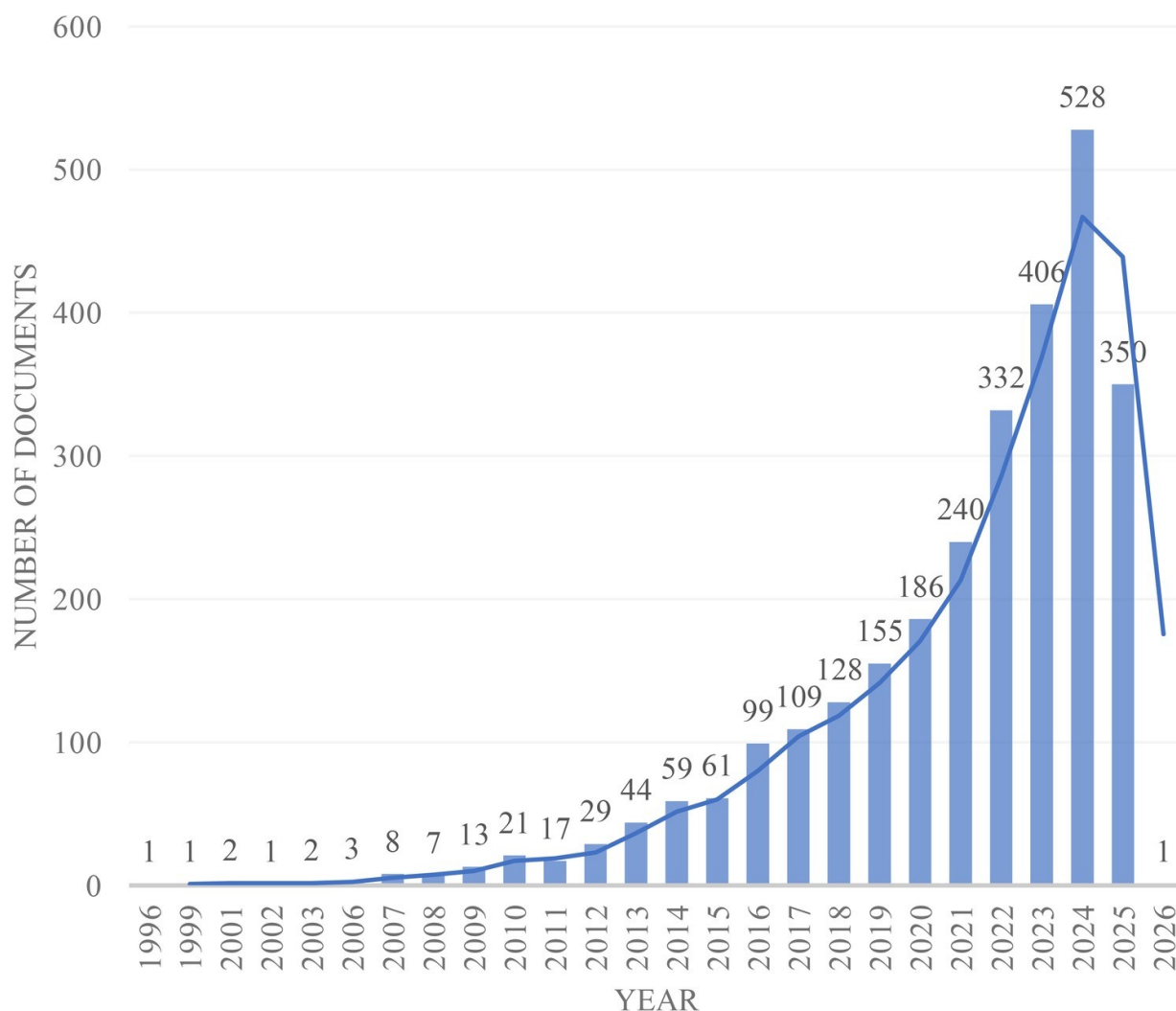


Figure 2. Number of publications on sentiment analysis in financial markets. (Source: Author's calculations based on Scopus results). [📄](#)

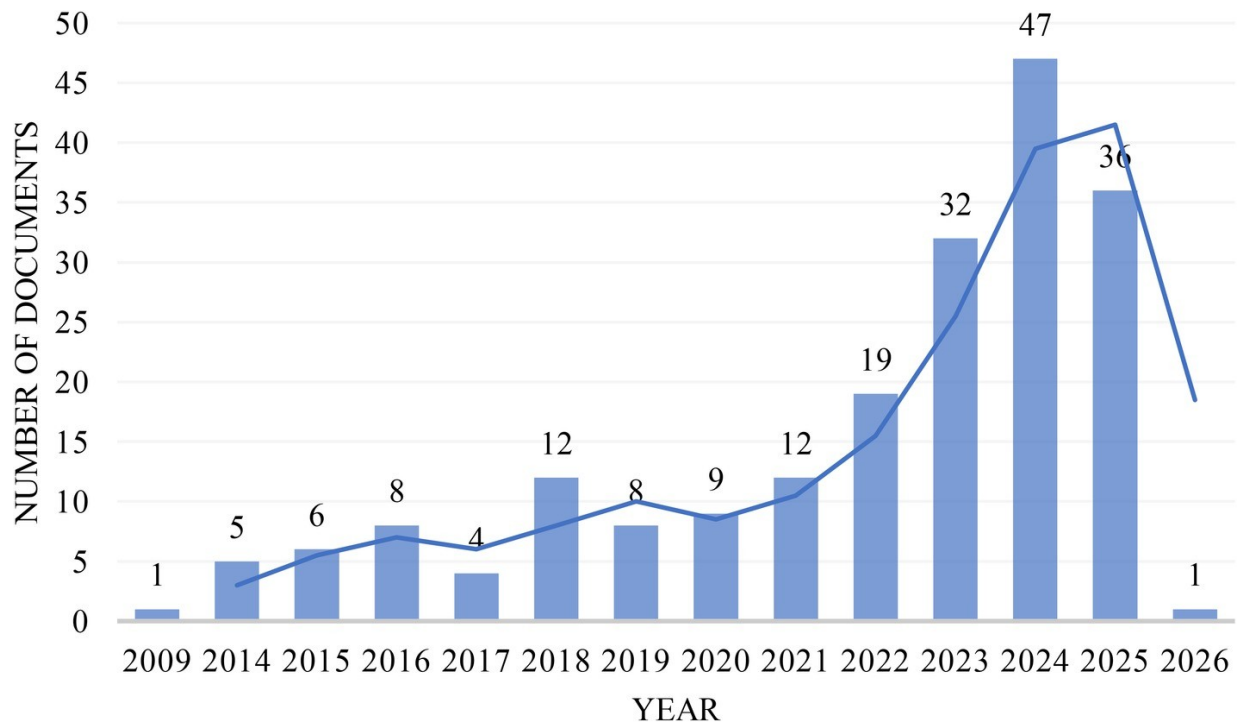


Figure 3. Number of Publications on Sentiment Analysis using AI tools in Stock Markets. (Source: Author's calculations based on Scopus results). [↩](#)

[Figure 3](#) is obtained by searching for the studies on sentiment analysis in the stock market context that have used AI applications. The rise in publications involving AI techniques for investor sentiment analysis is increasing rapidly indicating the growing importance of AI based tools in this domain. While the rise has been consistent year 2024 has witnessed 47 publications spread across multiple disciplines.

Based on the above findings, the year 2009 is considered the year of emergence of AI tools in investor sentiment analysis. 2009 is therefore the dividing year for Pre and Post AI applications in sentiment analysis in stock markets. Further, a mapping is done to study the spread of sentiment analysis across various domains and notable changes post the emergence of AI based applications.

4 STUDIES ACROSS DISCIPLINES

The field of sentiment analysis has attracted interest from scholars across various disciplines, with the highest number of publications in Computer Science accounting for 37 per cent, followed by Engineering (19 per cent), Mathematics (10 per cent), Decision Sciences (8 per cent), and Business, Management and Accounting (5 per

cent) (refer [Figure 4](#)) of the total documents published between 2009 and 31 July 31, which amounts to 200. When considering a broader period that includes all publications on sentiment analysis in financial markets ([Figure 5](#)), the publication pattern across discipline is drastically different with 28 per cent publications in computer science discipline and Economics, Econometrics and Finance has a share of 14 per cent, followed by Engineering (13 per cent), Business, Management and Accounting (10 per cent) and Mathematics (10 per cent). The share of Economics, Econometrics and Finance discipline is low with 4 per cent in studies where AI applications are adopted. Most of the studies in the economics and finance disciplines employ econometric tools and regression analysis and have different research objectives than AI applications-based studies.

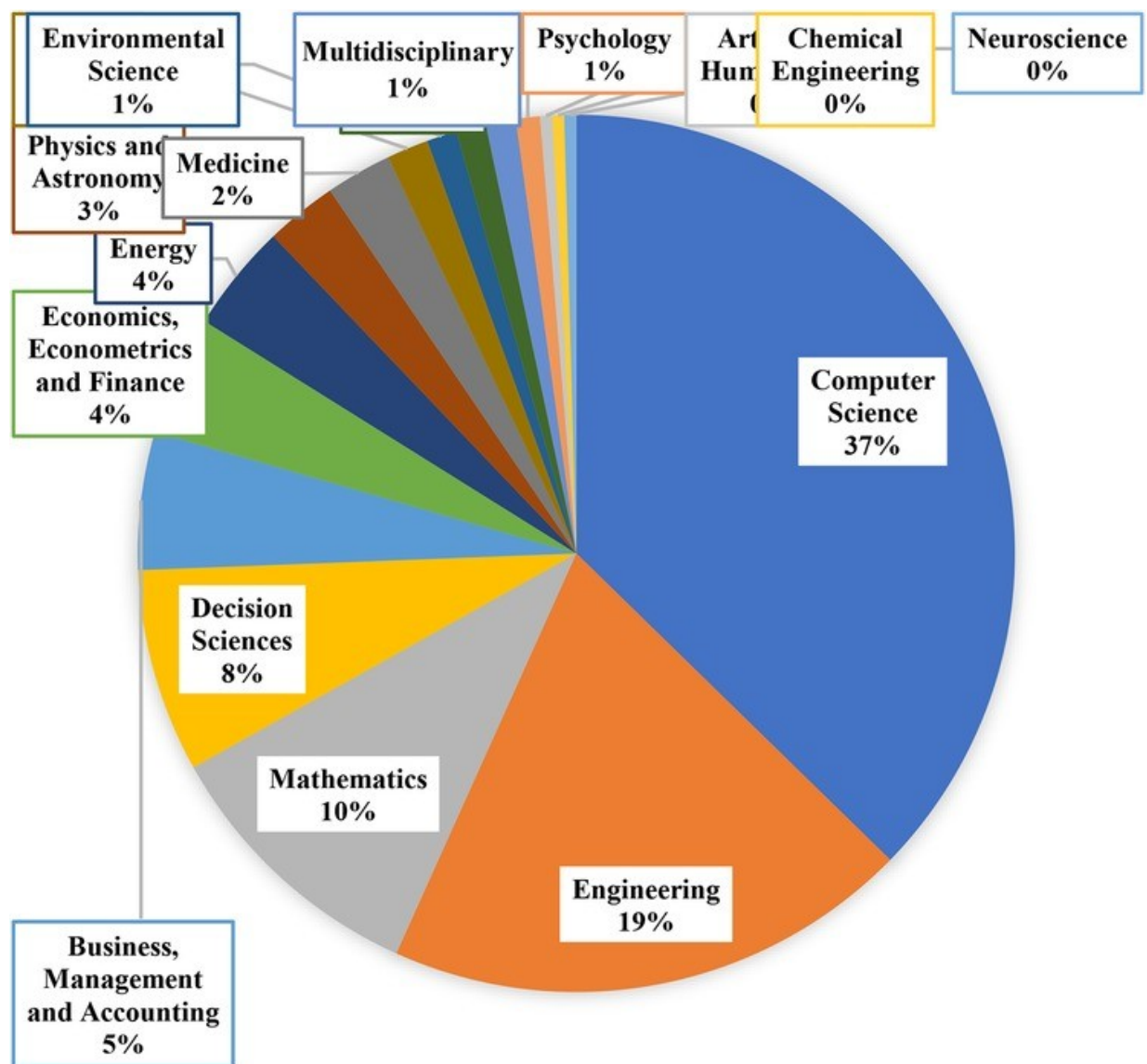


Figure 4. Documents by Discipline on Sentiment Analysis using AI tools in Financial Markets. [↗](#)

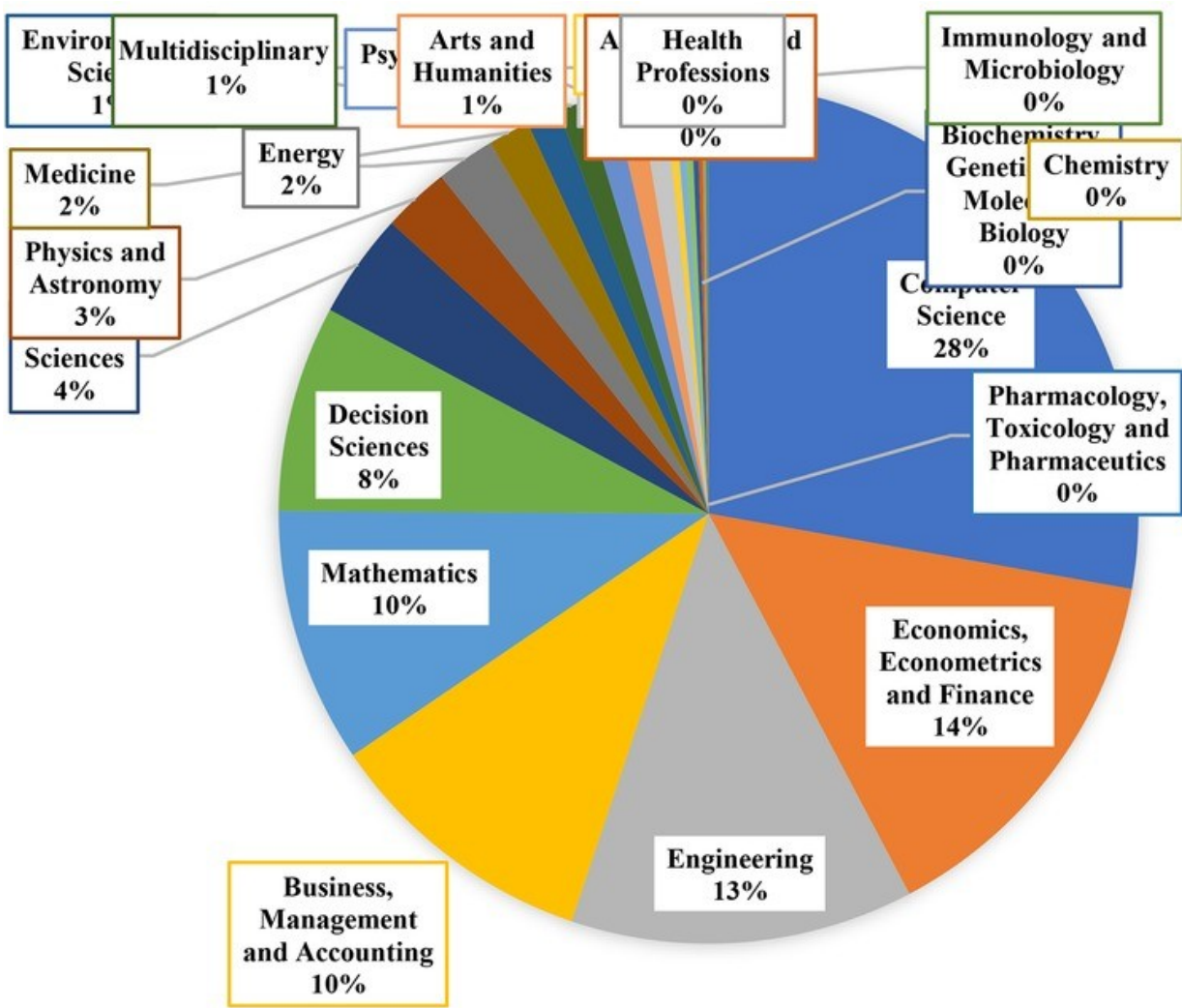


Figure 5. Documents by Discipline on Sentiment Analysis in Financial Markets. [↗](#)

5 KEYWORD ANALYSIS

Further, to understand the scholarly approach to sentiment analysis, keyword analysis is performed on the search results under (i) sentiment analysis in financial markets and (ii) sentiment analysis using AI tools in financial markets. Keyword analysis is useful for studying the knowledge structure and research segments in a domain([Lee and Su, 2010](#)). The authors' idea is explored through the analysis of the

keywords used by them. A keyword co-occurrence map is prepared based on keywords with a minimum of five occurrences using VOS viewer software.

5.1 *Sentiment analysis in financial markets*

[Table 1](#) shows that sentiment analysis has the highest number of occurrences, with the strongest link to other keywords, followed by machine learning with 333 occurrences, stock market with 333 occurrences, deep learning with 217 occurrences and LSTM with 136 occurrences. This analysis reveals that machine learning approaches are widely used by authors for sentiment analysis, particularly in the stock market, and many studies have used the LSTM (acronym for Long-Short term memory), natural language processing, text mining and random forest.

Table 1. Top 20 keywords on sentiment analysis in financial markets. [📄](#)

S. No.	Keyword	Occurrences
1	sentiment analysis	895
2	machine learning	333
3	stock market	333
4	deep learning	217
5	Lstm	136
6	investor sentiment	307
7	stock market prediction	139
8	Twitter	114
9	natural language processing	117
10	social media	115
11	stock prediction	100
12	stock price prediction	97
13	text mining	85
14	random forest	49
15	Sentiment	80
16	Prediction	52
17	Nlp	44
18	covid-19	81
19	data mining	45
20	big data	43

(Source: VOS viewer).

The keyword co-occurrence network in [Figure 6](#) illustrates the interconnection of different research clusters, where each node, represented by circles, indicates the keywords used by authors. The size of the node is directly proportional to the frequency of the keyword. Each edge, shown by the connecting line between nodes, indicates the co-occurrence between the connected keywords. The thickness of the edges reflects the frequency of co-occurrence of the connected keywords. In [Figure 6](#), Sentiment analysis is the most frequently used keyword connected to other clusters. It shows that studies related to sentiment analysis are linked to disciplines such as the stock market, ML, and investor sentiments.

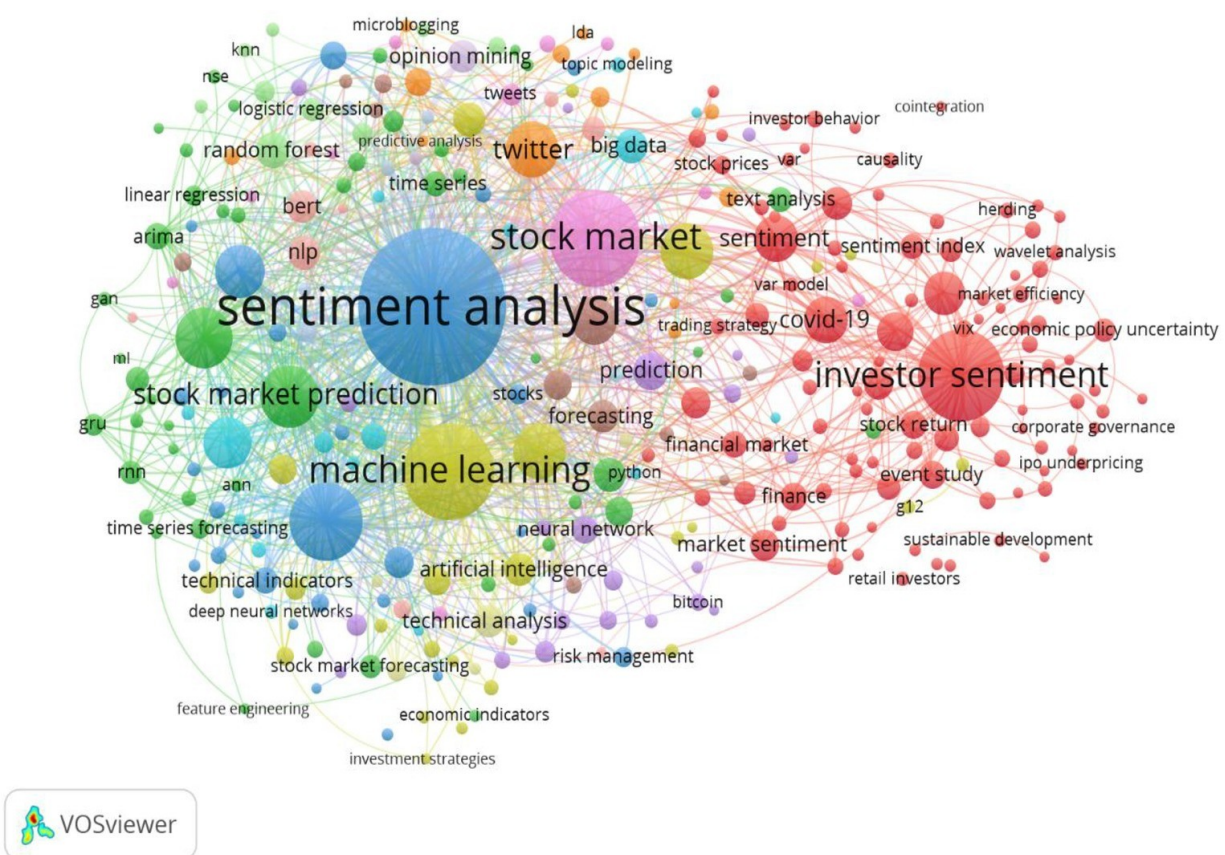


Figure 6. Keyword co-occurrence network based on 308 keywords with at least 5 occurrences. (Source: VOS viewer). [↗](#)

5.2 Sentiment analysis using AI tools in financial markets

Based the keyword analysis of 200 specific papers in Scopus on sentiment analysis using AI tools in financial markets as shown in [Table 2](#), the keyword with the highest frequency is sentiment analysis, followed by machine learning, artificial intelligence, stock market and deep learning. A specific search into the domain reveals that the scholarly work employing AI focuses on deep learning, natural language processing, stock market prediction, social media and artificial neural networks.

Table 2. Top 20 keywords on sentiment analysis using AI applications in financial markets 

S. No.	Keyword	Occurrences
1	sentiment analysis	81
2	machine learning	61
3	artificial intelligence	31
4	stock market	26
5	deep learning	25
6	natural language processing	20
7	stock market prediction	14
8	stock prediction	12
9	random forest	7
10	text mining	8
11	social media	10
12	artificial neural network	8
13	Twitter	8
14	Nlp	7
15	Lstm	7
16	Prediction	6
17	neural network	6
18	stock price prediction	7
19	artificial neural networks	6
20	big data	6

(Source: VOS viewer).

The keyword co-occurrence network in [Figure 7](#) shows that a segment of scholarly work is concentrated on applications of AI in sentiment analysis by employing ML techniques with a focus on deep learning models like artificial neural

networks for stock market prediction using social media data. This domain of research is connected to text mining and sentiment analysis. Natural language processing links methodological approaches like LSTM for stock price prediction. The diagram also shows an interdisciplinary research connecting finance with social media and AI.

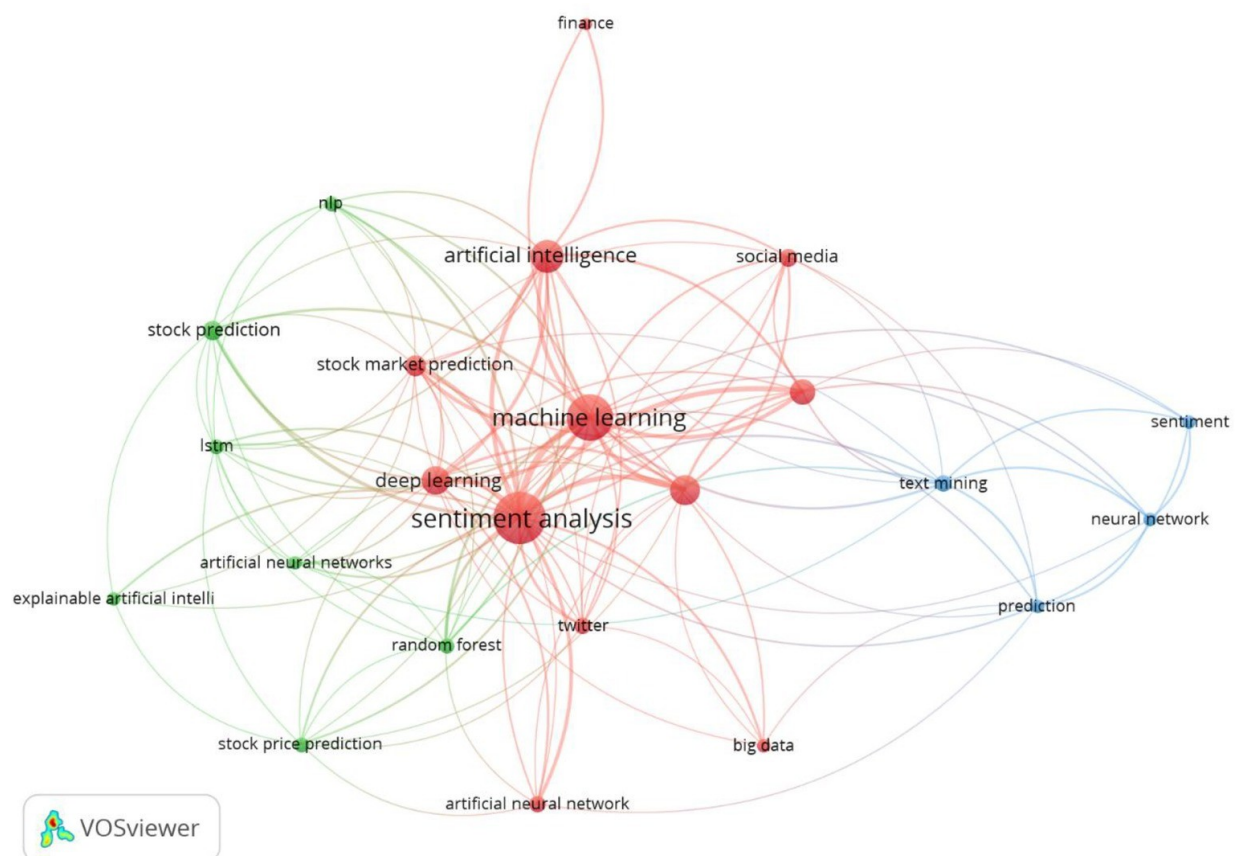


Figure 7. Keyword co-occurrence network based on 23 keywords with at least 5 occurrences. (Source: VOS viewer) [↗](#)

6 CONTRIBUTING JOURNALS ON SENTIMENT ANALYSIS USING AI APPLICATIONS IN FINANCIAL MARKETS

The study further locates the publication avenues that are widely contributing to the domain knowledge. These sources can be explored to map the knowledge structure, study the advances in the domain and identify unattended research issues. The distribution of articles on investor sentiments using AI applications in different journals is shown in figure below:

As evident from [Figure 8](#), journals in the Computer Science and finance disciplines are focusing on Articles on AI tools for sentiment analysis of individual investors.

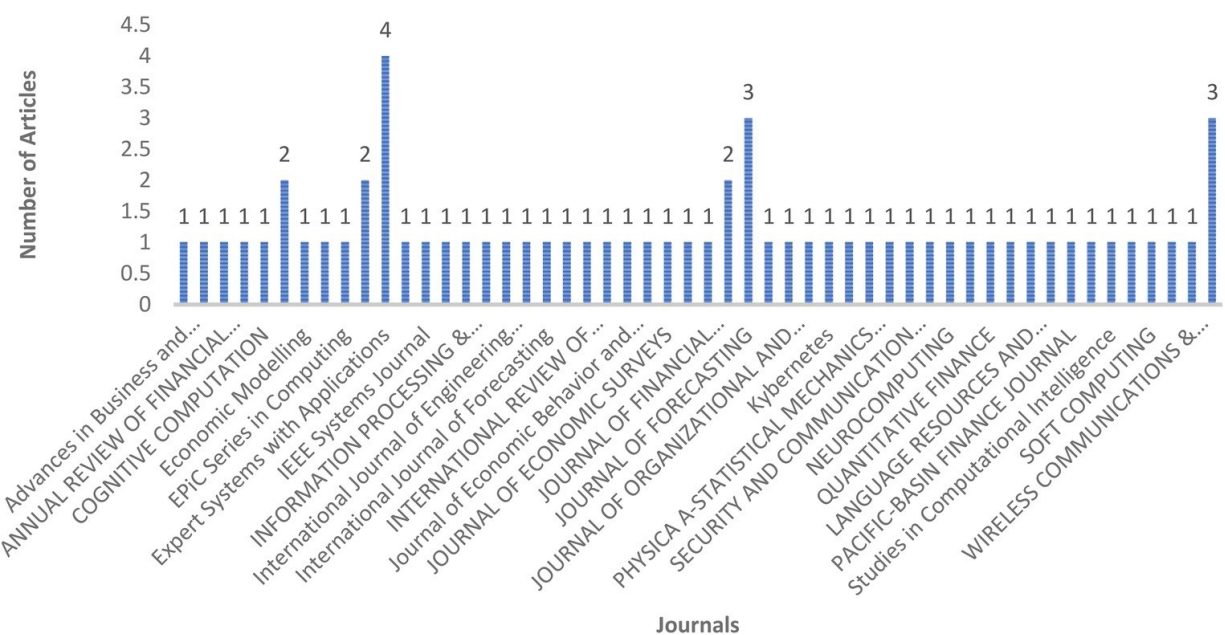


Figure 8. Journal-wise distribution of Articles. [↗](#)

(Source: Authors' compilation based on Scopus results).

7 COUNTRY COLLABORATION STRUCTURE

The study looks for the countries that publish and collaborate in this domain. The country collaboration map is prepared using VOS viewer software based on a minimum 5 publications of a country.

[Figure 9](#) shows that India has the highest number of publications followed by China and the United States. India collaborates with Italy, China and United States. China has collaboration with United Kingdom and Canada. The United States is closely working with Turkey. India is expanding its research network rapidly by collaborating with other nations and this network is a testament to the future potential of collaborative research with the connected countries.

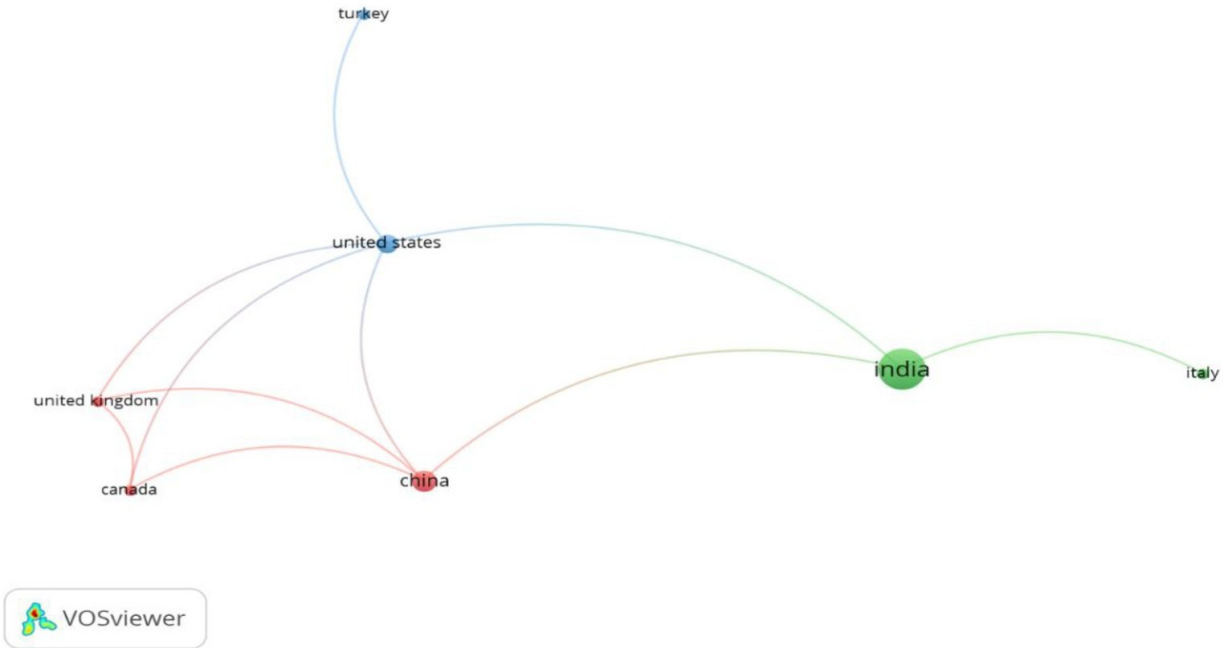


Figure 9. Country Collaboration Network based on at least 5 publications. (Source: VOS viewer). [↗](#)

8 CONTRIBUTING INSTITUTIONS

To further this field of investigation and map the landscape of Indian academia contributing to this domain, the Top 20 contributing institutions are identified based on the number of publications by institutions. Results in [Figure 10](#) show that Chandigarh University has highest number of publications, followed by Manipal Institute of Technology, Institute of Nano Science and Technology, and National Institute of Technology, Karnataka.

Study	Objectives	Sentiment	ML Algorithms/Deep learning Approaches	Dataset
<u>Yang <i>et al.</i>, (2018)</u>	Design an investor sentiment reward-based trading system.	Text	SVM, Gaussian process based inverse reinforcement learning algorithm	Thomson Reuters news analytics database
<u>Kim and Kim, 2014)</u>	Examine the relation between investor sentiment and future stock returns and constructing sentiment index.	Text	Naïve Bayes	Yahoo! Finance messages
<u>Wu <i>et al.</i>, (2014)</u>	Develop a novel sentiment ontology to conduct context-sensitive sentiment analysis of online opinion posts in stock markets.	Text	SVM, Generalized autoregressive conditional heteroskedasticity modeling.	Sina Finance

Study	Objectives	Sentiment	ML Algorithms/Deep learning Approaches	Dataset
<u>Van De Kauter et al., (2015)</u>	Investigate the use of sentiment analysis to detect positive and negative sentiment in financial news articles	Text	SVM, Coarse grained lexicon-based sentiment analysis	Financial News Articles
<u>Chan and Chong, (2017)</u>	To develop a language parser for sentiment analysis	Text	Decision tree, SVM	Movie reviews from Rotten Tomatoes website
Renault, (2017)	Study investor sentiment from messages posted on social media; and investigate the relationship between online investor sentiments and intra- day stock returns.	Text	Maximum entropy	StockTwist

Study	Objectives	Sentiment	ML Algorithms/Deep learning Approaches	Dataset
Huang and Zhang, (2022)	Study the relation between industry-level media tone and expected stock returns in China.	Text	Naive Bayes	A-shares listed on the Shanghai and Shenzhen Stock Exchanges
Oliveira <i>et al.</i> , (2017)	Assess the validity of microblogging data in forecasting of stock market variables.	Text	SVM, Random Forest , Ensemble Averaging (EA) method.	Twitter data
Zhang <i>et al.</i>, (2018)	Analyze investor's psychological characteristics and investment decision-making behaviour and examine sentiments in the network public opinion space.	Text	LSTM, CNN	Twitter data
Malandri <i>et al.</i> , (2018)	Examine the impact of mood on portfolio management.	Text	LSTM, multi-layer perceptron (MLP), and random forest classifier (RFC).	historical prices from NYSE

Study	Objectives	Sentiment	ML Algorithms/Deep learning Approaches	Dataset
Q. Wang et al., (2018)	Design a framework that combines the wisdom of crowds and technical analysis for financial market prediction using a new fusion strategy	Text	Deep random subspace ensembles	Sina Weibo and stock market records
Nam and Seong, (2019)	Forecast stock price movement with financial news considering causality.	Text	Transfer entropy, Multiple kernel learning	Naver
Ren et al., (2019)	Forecasting the direction of market movement.	Text	SVM	rReviews from Sina and Eastmoney forum
Ghosh et al., (2019)	To analyze and predict future growth of a company by stock price prediction	Text	LSTM, companies' net growth calculation algorithm	Twitter data

Study	Objectives	Sentiment	ML Algorithms/Deep learning Approaches	Dataset
Gurav and Kotrappa, (2020)	To learn short term and long term stock market 's sentiment pattern	Text	Log bilinear (LBL) RNN	Twitter data
Y. Li et al., (2020)	To predict investor sentiment	Text	LSTM, SVM and Naïve Bayes	stock message board of Eastmoney.com
Jin et al., (2020)	To predict closing price of stock	Text	LSTM	Facebook and Twitter
G. Wang <i>et al.</i> , (2021)	To investigate the relationships among environmental news, investor sentiment, and green industry stock returns in China.	Text	LSTM,Jieba	keywords in news headlines

Study	Objectives	Sentiment	ML Algorithms/Deep learning Approaches	Dataset
S. Li <i>et al.</i> , (2021)	Provide an interface for machine-learning and deep- learning (neural network) algorithms to acquire explainable sentiment-aware information from comment base	Text	LSTM, CNN, Gated Recurrent Unit (GRU), RNN.	comments on eastmoney.com
Obaid and Pukthuanthong, (2022)	To build a stock market level sentiment index	Picture	CNN	DeepSent data set
Issam <i>et al.</i>, (2022)	To create and test a system for prediction of tweets.	Text	CNN+LDA	Twitter

10 MACHINE LEARNING ALGORITHMS IN INVESTOR SENTIMENT ANALYSIS

The application of ML algorithms to study investor sentiments in markets and price predictability has gained tremendous attention of researchers in various domains. A number of deep learning based approaches are being applied to classify, process and analyze the sentiment of investors. ML based methods include Naïve Bayes, Support

Vector machine, Maximum entropy, Decision tree etc. it has been observed that supervised learning approaches are used more than the unsupervised learning. Various deep learning based approaches like Convolutional Neural Network (CNN), Recursive Neural Network (RvNN), Recurrent Neural Network (RNN), and Long Short Term Memory (LSTM) have also been applied to classify the sentiments.

11 DEEP LEARNING MODELS FOR INVESTOR SENTIMENT ANALYSIS

11.1 *Convolutional Neural Networks (CNN)*

CNNs offer the finest architecture for identifying and studying vital characteristics of images. They are highly used in the domains like Medical Imaging, Traffic Sign Detection, Speech Recognition, Face recognition etc. A CNN may consist of numerous layers, ranging from tens to hundreds, to identify various features within an image. Filters are utilized on each training image at multiple resolutions, with the output from each convolved image serving as the input for the subsequent layer. The filters can begin with very basic elements such as edges, curves, and brightness, gradually increasing in complexity to encompass the features that distinctly characterize the object. A CNN comprises an input layer, an output layer, and several intermediate layers. In any feedforward neural network, the intermediate layers are referred to as hidden layers, as their inputs and outputs are obscured by the activation function and the final convolution. As the convolutional kernel traverses the input matrix for the layer, the convolutional operation produces a feature map, which subsequently contributes to the input of the next layer. This process is succeeded by additional layers, including pooling layers, fully connected layers, and normalization layers.

[Zhang *et al.*, \(2018\)](#) used Long Short Term memory Convolutional network for sentiment classification. They analyzed the investor psychology to understand their sentiments. Obaid and Pukthuanthong (2022) used CNNs to study sentiment in photos uploaded by investors on various social media platforms. Similarly, You *et al.* (2015) used CNNs to analyze image sentiments on social media. S. Li *et al.* (2021) applied CNN to analyze investor sentiment on comments posted on eastmoney web portal.

11.2 *Recursive Neural Networks (RvNN)*

RvNNs process hierarchical data typically represented in a tree-like structure. It assigns the same set of weights to the structured input. They are useful in decoding sequences and making a tree structure in natural language processing. They use a structural representation of a sentence in the form of a parse tree. They can easily learn a hierarchical network. However, a major shortcoming of RvNNs is that each word's tag and a tree representation of the input sequence must be known. It is challenging for training purposes since the structure of each input sequence varies for every training sample.

11.3 *Recurrent Neural Networks (RNN)*

RNNs are used for modelling the sequential data. They are a variant of RvNNs, they take into account the processing time of sequential elements. In order to anticipate the next word in a phrase, RNNs process each word repeatedly, storing internal states of information. When a network classifies a text passage into several classification tasks, RNNs are a frequent choice for sentiment classification. One of the most popular RNN variations is LSTM. Various studies in the past have used LSTM for sentiment classification.

11.4 *Long Short Term Memory (LSTM)*

LSTMs were developed to overcome the vanishing gradient problem while training the RNNs. Scholar have used LSTM in predicting sentiments of investors in the market. Malandri *et al.* (2018) studied how public mood affects portfolio managements. They applied LSTM, MLP, and RFC to generate the portfolios. Huang *et al.* (2017) used CNN and LSTM and proposed an architecture with one layer of CNN and two layers of LSTM stacked on CNN. Similarly, [Wang *et al.* \(2021a\)](#) applied LSTM to predict investor sentiment for investigating the relationship between environmental news, investors' sentiments and stock returns in China. The data was taken Jieba and Word2Vec datasets.

12 MACHINE LEARNING ALGORITHMS

12.1 *Naïve Bayes classification algorithm*

Naïve Bayes algorithm is a popular and straightforward NLP algorithm. The Naïve Bayes categorization divides a text message into word groups. Every text message

has a single label, and every word is a feature. The probability of a label given features is calculated using the Bayes rule as follows:

$$P(\text{label} / \text{features}) = [P(\text{label}) * P(\text{features} / \text{label})] / P(\text{features})$$

Antweiler and Frank (2004) examined the accuracy of Naïve Bayes Classification. They manually categorized 1000 messages and applied the algorithm to verify the accuracy and an accuracy of 72.3% was achieved. Das and Chen (2007) reported an approximate 50% classification accuracy of 50% on Yahoo! Finance messages.

12.2 *Support Vector Machine*

Support vector machine can solve nonlinear problems by transforming them to a quadratic programming ([Ren et al., 2019b](#)). The SVM model was introduced for classification tasks. Later, it adopted an ϵ -insensitive loss function and expanded to regression ([Oliveira et al., 2017b](#)). By choosing the greatest margin hyperplane in feature space, the widely used SVM approach can lessen the issue of overfitting ([Yu et al., 2009](#)).

12.3 *Maximum entropy Classifier*

It is a probabilistic classifier that belongs to the class of exponential models. It is useful for text classification and highly used in sentiment analysis. Renault, (2017) used maximum entropy classifier to compare the out-of-sample classification accuracy.

13 MAJOR DATABASES

StockTwits: It is a popular social media platform for investors to share stock and financial market-related content. This platform is widely used by researchers for studying investor sentiments. Renault (2017) constructed a lexicon of words used by online investors when they share their opinions on microblogging platform StockTwits.

DeepSent dataset: This dataset is used for sentiment classification in images. The dataset contains 1269 images classified as positive or negative. It allows binary sentiment classification of images. Obaid and Pukthuanthong (2022) used DeepSent dataset for training Google Inception V3 model.

Eastmoney: Eastmoney is an investment related platform in China used by investors for trading stocks listed in Chinese stock exchanges (S. Li *et al.*, 2021). Ren *et al.* (2019) used eastmoney for accessing investor comments.

X, formerly Twitter: Various other scholars relied on X formerly Twitter and other social media platform for gathering sentiment data. Oliveira *et al.*, (2017) used Twitter data to test its validity for forecasting stock market variables.

14 CONCLUSION

This study highlights the significance of investor sentiment analysis for market stability. It shows the growth of sentiment analysis and application of AI tools in sentiment analysis in financial markets. Using bibliometric and systematic literature review, the paper presents publication trends, knowledge structure and collaboration network. A review of major contributions in the field of sentiment analysis is performed to study the advances in the domain of sentiment analysis in financial markets. It is found that LSTM performs better for sentiment analysis in text and CNNs have also made a significant contribution in image analysis. The study also discusses the datasets used by the researchers for collecting sentiment data or to train the models. The findings are useful for regulators to monitor financial and stock related discussions from user accounts and take appropriate policy measures. Further, it is suggested that future studies should focus on analyzing emotions, voice and video data and also multimodal data should also be used for analyzing sentiments.

REFERENCES

- Baig, A.S., Blau, B.M., Butt, H.A., and Yasin, A. (2022). Do retail traders destabilize financial markets? An investigation surrounding the COVID-19 pandemic. *Journal of Banking and Finance*, 144, 106627. <https://doi.org/10.1016/j.jbankfin.2022.106627>
- Baker, M., and Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, 21(2), 129–152. <https://doi.org/10.1257/JEP.21.2.129>
- Barber, B.M., and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21(2), 785–818. <https://doi.org/10.1093/rfs/hhm079>

- Chan, S.W.K., and Chong, M.W.C. (2017). Sentiment analysis in financial texts. *Decision Support Systems*, 94, 53–64. <https://doi.org/10.1016/j.dss.2016.10.006>
- Cheng, F., Chiao, C., Wang, C., Fang, Z., and Yao, S. (2021). Does retail investor attention improve stock liquidity? A dynamic perspective. *Economic Modelling*, 94, 170–183. <https://doi.org/10.1016/j.econmod.2020.10.001>
- Ghosh, A., Bose, S., Maji, G., Debnath, N.C., and Sen, S. (2019). Stock price prediction using lstm on indian share market. *EPiC Series in Computing*, 63, 101–110. <https://doi.org/10.29007/QGCZ>
- Goetzmann, W.N., and Kumar, A. (2008). Equity portfolio diversification. *Review of Finance*, 12(3), 433–463. <https://doi.org/10.1093/rof/rfn005>
- Gurav, U., and Kotrappa, S. (2020). Predict stock market's fluctuating behaviour : Role of investor's sentiments on stock market performance. *International Journal of Engineering Trends and Technology*, 68, 72–80. <https://doi.org/10.14445/22315381/IJETT-V68I11P209>
- Han, B., and Kumar, A. (2013). Speculative retail trading and asset prices. *Journal of Financial and Quantitative Analysis*, 48(2), 377–404. <https://EconPapers.repec.org/RePEc:cup:jfinqa:v:48:y:2013:i:02:p:377-404>
- Haritha P. H. (2024). The effect of heuristics on indian stock market investors: Investor sentiment as a mediator. *Management and Labour Studies*, 49(1), 43–61. <https://doi.org/10.1177/0258042X231170745>
- Huang, T., and Zhang, X. (2022). Industry-level media tone and the cross-section of stock returns. *International Review of Economics and Finance*, 77, 59–77. <https://doi.org/10.1016/J.IREF.2021.09.002>
- Issam, A., Mounir, A.K., Saida, E.M., and Fatna, E.M. (2022). Financial sentiment analysis of tweets based on deep learning approach. *Indonesian Journal of Electrical Engineering and Computer Science*, 25(3), 1759–1770. <https://doi.org/10.11591/IJEECS.V25.I3.PP17591770>
- Jin, Z., Yang, Y., and Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. *Neural Computing and Applications*, 32(13), 9713–9729. <https://doi.org/10.1007/S00521-019-04504-2/METRICS>
- Kim, S.H., and Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior and Organization*, 107(PB), 708–729. <https://doi.org/10.1016/J.JEBO.2014.04.015>
- Lee, P.C., and Su, H.N.N. (2010). Investigating the structure of regional innovation system research through keyword co-occurrence and social network analysis. *Innovation*, 12(1), 26–40. <https://doi.org/10.5172/impp.12.1.26>
- Li, S., Ning, K., and Zhang, T. (2021a). Sentiment-aware jump forecasting. *Knowledge-Based Systems*, 228, 107292. <https://doi.org/10.1016/J.KNOSYS.2021.107292>

- Li, S., Ning, K., and Zhang, T. (2021b). Sentiment-aware jump forecasting. *Knowledge-Based Systems*, 228. <https://doi.org/10.1016/j.knosys.2021.107292>
- Li, Y., Bu, H., Li, J., and Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning. *International Journal of Forecasting*, 36(4), 1541–1562. <https://doi.org/10.1016/J.IJFORECAST.2020.05.001>
- Linnainmaa, J. (n.d.). Who Makes the Limit Order Book? Implications for Contrarian Strategies, Attention-Grabbing Hypothesis, and the Disposition Effect.
- Malandri, L., Xing, F.Z., Orsenigo, C., Vercellis, C., and Cambria, E. (2018a). Public mood–driven asset allocation: the importance of financial sentiment in portfolio management. *Cognitive Computation*, 10(6), 1167–1176. <https://doi.org/10.1007/S12559-018-9609-2/METRICS>
- Malandri, L., Xing, F.Z., Orsenigo, C., Vercellis, C., and Cambria, E. (2018b). Public mood–driven asset allocation: the importance of financial sentiment in portfolio management. *Cognitive Computation*, 10(6), 1167–1176. <https://doi.org/10.1007/s12559-018-9609-2>
- Mechanics and Its Applications, 573. <https://doi.org/10.1016/j.physa.2021.125979>
- Nam, K.H., and Seong, N.Y. (2019). Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. *Decision Support Systems*, 117, 100–112. <https://doi.org/10.1016/J.DSS.2018.11.004>
- Obaid, K., and Pukthuanthong, K. (2022a). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics*, 144(1), 273–297. <https://doi.org/10.1016/J.JFINECO.2021.06.002>
- Obaid, K., and Pukthuanthong, K. (2022b). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics*, 144(1), 273–297. <https://doi.org/10.1016/j.jfineco.2021.06.002>
- Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance*, 53(5), 1775–1798. <https://doi.org/10.1111/0022-1082.00072>
- Odean, T. (1999). Do investors trade too much? *American Economic Review*, 89(5), 1279–1298. <https://doi.org/10.1257/AER.89.5.1279>
- Oliveira, N., Cortez, P., and Areal, N. (2017a). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications*, 73, 125–144. <https://doi.org/10.1016/J.ESWA.2016.12.036>
- Oliveira, N., Cortez, P., and Areal, N. (2017b). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading

- volume and survey sentiment indices. *Expert Systems with Applications*, 73, 125–144. <https://doi.org/10.1016/j.eswa.2016.12.036>
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/BMJ.N71>
- Ren, R., Wu, D.D., and Wu, D.D. (2019a). Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Systems Journal*, 13(1), 760–770. <https://doi.org/10.1109/JSYST.2018.2794462>
- Ren, R., Wu, D.D., and Wu, D.D. (2019b). Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Systems Journal*, 13(1), 760–770. <https://doi.org/10.1109/JSYST.2018.2794462>
- Renault, T. (2017a). Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking and Finance*, 84, 25–40. <https://doi.org/10.1016/J.JBANKFIN.2017.07.002>
- Renault, T. (2017b). Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking and Finance*, 84, 25–40. <https://doi.org/10.1016/j.jbankfin.2017.07.002>
- Schumaker, R.P., Zhang, Y., Huang, C.N., and Chen, H. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems*, 53(3), 458–464. <https://doi.org/10.1016/j.dss.2012.03.001>
- Shleifer, A., and Summers, L.H. (1990). The Noise Trader Approach to Finance. In *Journal of Economic Perspectives* (Vol. 4).

Snow, N.M., and Rasso, J. (2025). If the Tweet Fits: How Investors Process Financial Information Received via Social Media. *Journal of Information Systems*, 39(2), 93–110. <https://doi.org/10.2308/ISYS-2022-044>

Van De Kauter, M., Breesch, D., and Hoste, V. (2015). Fine-grained analysis of explicit and implicit sentiment in financial news articles. *Expert Systems with Applications*, 42(11), 4999–5010. <https://doi.org/10.1016/J.ESWA.2015.02.007>

Wang, G., Yu, G., and Shen, X. (2021a). The effect of online environmental news on green industry stocks: The mediating role of investor sentiment. *Physica A: Statistical Mechanics and Its Applications*, 573, 125979. <https://doi.org/10.1016/J.PHYSA.2021.125979>

Wang, G., Yu, G., and Shen, X. (2021b). The effect of online environmental news on green industry stocks: The mediating role of investor sentiment. *Physica A: Statistical Mechanics and Its Applications*, 573. <https://doi.org/10.1016/j.physa.2021.125979>

- Wang, Q., Xu, W., and Zheng, H. (2018). Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles. *Neurocomputing*, 299, 51–61.
<https://doi.org/10.1016/J.NEUCOM.2018.02.095>
- Wu, D.D., Zheng, L., and Olson, D.L. (2014). A decision support approach for online stock forum sentiment analysis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(8), 1077–1087.
<https://doi.org/10.1109/TSMC.2013.2295353>
- Yang, S.Y., Yu, Y., and Almahdi, S. (2018). An investor sentiment reward-based trading system using Gaussian inverse reinforcement learning algorithm. *Expert Systems with Applications*, 114, 388–401.
<https://doi.org/10.1016/J.ESWA.2018.07.056>
- Yu, L., Chen, H., Wang, S., and Keung Lai, K. (n.d.). Evolving Least Squares Support Vector Machines for Stock Market Trend Mining *.
- Yu, L., Chen, H., Wang, S., and Lai, K.K. (2009). Evolving least squares support vector machines for stock market trend mining. *IEEE Transactions on Evolutionary Computation*, 13(1), 87–102.
<https://doi.org/10.1109/TEVC.2008.928176>
- Zhang, G., Xu, L., and Wang, L. (2018). Sentiments classification in stock network public opinion space based on long-short memory convolution neural network. *MATEC Web of Conferences*, 189.
<https://doi.org/10.1051/MATECCONF/201818910010>

Artificial Intelligence and Ecology: A Humane Framework for Mapping Climate Futures in Canada and India

Kumar Ram Krishna

University of Delhi

Mahmoud Maseili

Formerly with the University of Ottawa, Canada

DOI: [10.1201/9781042018116-7](https://doi.org/10.1201/9781042018116-7)

ABSTRACT: The intensifying impacts of climate change have underscored the need for analytical tools capable of capturing the complexity, scale, and uncertainty of ecological transformation. Artificial intelligence (AI) has emerged as a critical enabler in this context, offering advanced capabilities for processing large environmental datasets, modelling non-linear climate systems, and generating forward-looking scenarios. This paper advances a humane framework for integrating AI into ecological analysis, arguing that the value of AI in mapping climate futures lies not only in predictive accuracy but also in its alignment with ethical responsibility, social equity, and ecological care. Drawing on examples from climate modelling, remote sensing, biodiversity monitoring, and early-warning systems, the study examines how AI can enhance anticipatory governance and adaptive decision-making while remaining accountable to human and non-human vulnerabilities. The paper critically engages with concerns surrounding data bias, energy-intensive computation, and techno-solutionism, emphasizing the need for transparent, context-sensitive, and participatory AI systems. By situating AI as a decision-support architecture rather than an autonomous authority, the article demonstrates how humane, human-centred AI can contribute to more just and sustainable climate futures.

Keywords: Ethical AI, Climate Change Modelling, Ecological Sustainability, Anticipatory Governance, and Human-Centred AI

1 METHODOLOGY

This study adopts a qualitative, interdisciplinary, and comparative methodology that integrates climate science literature, ecological history, and critical AI studies to examine the role of artificial intelligence in mapping climate futures in India and Canada. The analysis synthesizes secondary data from authoritative sources—including IPCC assessments, national climate reports, and peer-reviewed research—alongside conceptual insights from science and technology studies (STS), environmental ethics, and humane AI scholarship. Rather than conducting primary model development, the chapter employs interpretive analysis to examine how AI-enabled climate mapping tools (such as machine learning–based climate modelling, remote sensing, and early-warning systems) are framed, deployed, and governed within differing socio-ecological contexts. A comparative lens is used to identify asymmetries in vulnerability, governance capacity, and ethical implications across tropical and high-latitude systems, while a historical perspective (c. 1800–2025) situates contemporary AI applications within longer trajectories of ecological change. This methodological approach prioritizes ethical reflexivity, contextual sensitivity, and normative evaluation, positioning AI as a decision-support architecture embedded within human and ecological responsibility rather than as a neutral or autonomous predictive authority.

2 INTRODUCTION

The deepening climate crisis has transformed ecology into a domain defined by uncertainty, non-linearity, and cascading risk. Accelerating temperature rise, extreme weather events, biodiversity loss, and ecological tipping points now interact across temporal and spatial scales that exceed the capacity of conventional analytical methods to fully capture or anticipate ([IPCC, 2023](#)). As climate change increasingly manifests as a systemic challenge rather than a discrete environmental problem, there is a growing need for tools capable of integrating vast datasets, modelling complex feedback loops, and supporting anticipatory governance. Within this context, artificial intelligence (AI) has emerged as a significant analytical resource for ecological research and climate futures mapping ([Reichstein *et al.*, 2019](#)).

AI-driven techniques—ranging from machine learning–based climate models and satellite image analysis to biodiversity monitoring and disaster early-warning systems—are reshaping how ecological change is observed, interpreted, and projected ([Rolnick *et al.*, 2019](#)). By enabling the synthesis of heterogeneous data

sources such as remote sensing, meteorological records, oceanic observations, and socio-environmental indicators, AI enhances the resolution and predictive capacity of climate analysis. However, the growing prominence of AI in environmental governance also raises critical ethical, political, and ecological questions. Technological efficiency alone cannot address issues of environmental justice, unequal vulnerability, data asymmetry, or the social consequences of algorithmic decision-making ([Jasanoff, 2016](#)).

This chapter advances a humane framework for understanding the relationship between artificial intelligence and ecology. A humane perspective treats AI not as a technocratic solution to climate change, but as a decision-support system embedded within human values, institutional constraints, and ecological responsibilities. Such a framework emphasizes transparency, accountability, and inclusivity, while recognizing the intrinsic value of both human and non-human life systems (Floridi *et al.*, 2018). Rather than replacing human judgment, AI is positioned as an interpretive and enabling architecture—one that can assist in mapping climate futures, identifying early warning signals, and supporting adaptive responses, while remaining subject to ethical oversight and contextual understanding.

By situating AI within a broader ecological and normative landscape, this chapter argues that the future relevance of artificial intelligence in climate governance will depend not only on its computational power but on how responsibly it is designed, deployed, and integrated into climate decision-making processes. In doing so, it seeks to bridge the gap between technological innovation and ecological ethics, offering a framework through which AI can contribute meaningfully to more just, sustainable, and humane climate futures.

3 CLIMATE CHANGE IN INDIA AND CANADA: ECOLOGICAL IMPACTS AND THE ROLE OF ARTIFICIAL INTELLIGENCE IN MAPPING CLIMATE FUTURES

Climate change has emerged as one of the most consequential drivers of ecological transformation in both India and Canada, albeit through regionally distinct pathways shaped by geography, latitude, and socio-ecological vulnerability. Globally, mean surface temperatures have risen by approximately 1.2 °C above pre-industrial (1850–1900) levels, with observable impacts on atmospheric circulation, cryospheric stability, biodiversity, and hydrological systems ([IPCC, 2023](#)). India and Canada represent contrasting ecological theatres within this global crisis: India as a monsoon-dependent, densely populated tropical–subtropical system, and Canada as a high-latitude, cryosphere-sensitive ecological domain.

In India, instrumental records indicate a statistically significant warming trend since the early twentieth century, with mean surface temperatures increasing by approximately 0.7 °C between 1901 and 2018 ([Ministry of Earth Sciences \[MoES\], 2020](#)). This warming has been accompanied by increased frequency and intensity of heatwaves, heightened variability of the southwest monsoon, and a rise in short-duration extreme rainfall events, contributing to floods, droughts, and agricultural stress ([Krishnan *et al.*, 2020](#)). Ecologically, these changes have disrupted phenological cycles, accelerated Himalayan glacier retreat, and intensified pressure on freshwater systems and coastal ecosystems, particularly in the Indo-Gangetic plain and eastern seaboard (IPCC, 2022). The long historical baseline suggests relatively stable climatic conditions during the early nineteenth century, with pronounced acceleration in warming occurring after the mid-twentieth century, consistent with global anthropogenic forcing trends.

Canada, by contrast, has experienced warming at a rate exceeding the global average, particularly in northern and interior regions. National assessments show that Canada's mean temperature increased by approximately 1.7 °C between 1948 and 2023, with Arctic and sub-Arctic regions warming at more than twice the global rate ([Environment and Climate Change Canada \[ECCC\], 2024](#)). The ecological consequences include widespread permafrost thaw, declining sea ice extent, increased wildfire frequency, and shifts in boreal and tundra ecosystems ([Bush and Lemmen, 2019](#)). Long-term reconstructions suggest that pre-industrial Canadian climates were substantially colder, with recent warming representing a sharp departure from historical variability.

In both regions, the scale and complexity of ecological change have outpaced the capacity of traditional linear analytical methods. Artificial intelligence (AI) has therefore emerged as a critical methodological and governance tool for mapping climate futures. AI-driven machine learning models enable the integration of heterogeneous datasets—satellite imagery, climate model outputs, hydrological data, and ecological indicators—to identify non-linear patterns, feedback loops, and early warning signals ([Reichstein *et al.*, 2019](#)). In India, AI applications are increasingly used to model monsoon behaviour, project glacier melt in the Himalayas, assess coastal vulnerability, and support disaster early-warning systems, enhancing adaptive capacity in climate-sensitive regions ([Rolnick *et al.*, 2019](#); [MoES, 2020](#)). In Canada, AI supports high-resolution climate downscaling, wildfire prediction, urban heat mapping, and biodiversity monitoring, particularly in Arctic and boreal ecosystems where observational data are sparse ([ECCC, 2024](#)).

Crucially, AI's contribution to ecological mapping is not merely computational but interpretive. When embedded within a humane framework, AI functions as a decision-support architecture that augments human judgment rather than replacing

it, enabling more equitable, transparent, and ecologically sensitive climate governance. By revealing long-term trends, regional asymmetries, and future risk pathways, AI allows policymakers and researchers in both India and Canada to move from reactive responses toward anticipatory and adaptive ecological planning.

Table 1. Long-Term Climate Indicators for India and Canada (1800–2025).

Period	India: Temperature Trend	Canada: Temperature Trend	Key Ecological Signals
1800–1850	Pre-industrial baseline	Pre-industrial baseline	Stable monsoon patterns; extensive cryosphere
1851–1900	Minimal warming	Slight warming	Early industrial influence
1901–1950	+0.2–0.3 °C	+0.3–0.5 °C	Initial glacier retreat; ecosystem stability largely intact
1951–2000	+0.4–0.6 °C	+0.8–1.2 °C	Accelerated warming; Arctic amplification begins
2001–2025	+0.9–1.2 °C	+1.5–2.5 °C (regional)	Heatwaves (India); permafrost thaw and wildfires (Canada)

Note: Values are approximate ranges derived from national assessments and IPCC syntheses; early-period estimates rely on proxy reconstructions.

4 HOW CLIMATE CHANGE IS BECOMING A HUMANE CRISIS: SPECIES LOSS, GLACIAL RETREAT, AND HUMAN VULNERABILITY IN INDIA AND CANADA (1800–2025)

Climate change has progressively transformed from a geophysical phenomenon into a profoundly humane crisis, reshaping the conditions under which both human and non-human life systems survive. Over the past two centuries, anthropogenic warming has disrupted ecological equilibria that evolved over millennia, accelerating species extinction, destabilizing cryospheric systems, and intensifying social vulnerability. While the drivers of climate change are global, its humane consequences are regionally differentiated. India and Canada—representing densely populated tropical systems and sparsely populated high-latitude ecologies

respectively—illustrate how climate change simultaneously erodes ecological resilience and human security across contrasting environmental contexts ([IPCC, 2023](#)).

Since the early nineteenth century, biodiversity loss has accelerated at rates unprecedented in the geological record. Global species extinction rates are now estimated to be tens to hundreds of times higher than background levels, driven primarily by climate change interacting with habitat destruction and resource exploitation (IPBES, 2019). In India, rising temperatures and altered precipitation patterns have disrupted ecological niches across forests, wetlands, and coastal zones. Climate stress has contributed to range contractions among endemic species, coral bleaching in the Lakshadweep and Andaman regions, and declining pollinator populations critical to agrarian livelihoods ([Krishnan *et al.*, 2020](#)). These losses are not abstract ecological indicators; they undermine food security, medicinal knowledge systems, and cultural relationships with nature.

In Canada, climate-induced biodiversity stress is particularly evident in Arctic and boreal ecosystems. Rapid warming has altered tundra vegetation, enabling shrub expansion at the expense of cold-adapted species, while marine warming and sea-ice decline threaten keystone species such as polar bears, seals, and Arctic fish stocks ([Bush and Lemmen, 2019](#)). Over the past two centuries, ecosystems that once changed gradually are now being reorganized within decades, compressing evolutionary adaptation timelines beyond biological capacity. Species loss thus becomes a humane issue insofar as it dismantles ecological services upon which Indigenous communities, rural populations, and future generations depend.

Glaciers, long regarded as stable reservoirs of freshwater and climatic memory, have emerged as sensitive indicators of anthropogenic warming. Since the mid-nineteenth century, global glacier mass has declined steadily, with accelerated losses after 1950 (IPCC, 2022). In the Indian Himalayan region, glaciers that sustained perennial river systems such as the Indus, Ganga, and Brahmaputra are retreating at increasing rates, threatening long-term water availability for hundreds of millions of people ([Krishnan *et al.*, 2020](#)). Historically, these glaciers exhibited relative stability during the early nineteenth century; their rapid shrinkage in the late twentieth and early twenty-first centuries represents a sharp rupture from historical climatic norms.

Canada's cryosphere has experienced even more pronounced transformation. Mountain glaciers in western Canada and extensive Arctic ice systems have lost significant mass since the late nineteenth century, with northern regions warming at more than twice the global average ([Bush and Lemmen, 2019](#)). Permafrost thaw has destabilized infrastructure, altered hydrological systems, and released greenhouse gases, creating feedback loops that further accelerate warming. The humane

dimension of glacial loss lies in its temporal injustice: communities inherit risks generated by historical emissions they did not produce, while future generations face irreversible hydrological uncertainty.

Over the past two centuries, climate change has increasingly intersected with social inequality, governance capacity, and demographic pressure. In India, climate impacts manifest through intensifying heatwaves, erratic monsoons, floods, and droughts—events that disproportionately affect the poor, elderly, and informal labor populations (IPCC, 2022). Mortality from extreme heat and climate-related disasters has risen sharply since the late twentieth century, turning climate variability into a public health emergency. These impacts represent a departure from nineteenth-century climatic stability, when seasonal predictability underpinned agrarian and social systems.

In Canada, although economic capacity is higher, climate change is reshaping human security in subtler but profound ways. Indigenous and northern communities face food insecurity due to disrupted hunting ecosystems, unsafe ice conditions, and shifting species distributions ([Bush and Lemmen, 2019](#)). Urban populations increasingly confront heat stress and wildfire smoke, while climate-related disasters impose escalating economic and psychological costs. Across both regions, climate change erodes not only material well-being but also cultural continuity and intergenerational trust.

Artificial intelligence has emerged as a critical analytical instrument for understanding climate change as a humane phenomenon rather than merely a physical one. AI enables the integration of long-term climate reconstructions, satellite observations, ecological indicators, and socio-economic data to detect patterns of vulnerability and ecological stress that unfold across centuries ([Reichstein *et al.*, 2019](#)). Machine learning models help reconstruct historical climate baselines, identify thresholds of ecosystem collapse, and project future risk trajectories for both human and non-human systems.

In India, AI supports glacier monitoring, monsoon forecasting, and disaster early-warning systems, helping translate climatic signals into actionable humanitarian responses ([Rolnick *et al.*, 2019](#)). In Canada, AI-driven models assist in wildfire prediction, permafrost monitoring, and Arctic ecosystem mapping, enhancing adaptive planning in data-sparse regions ([Bush and Lemmen, 2019](#)). Importantly, when embedded within a humane framework, AI does not replace ethical judgment but strengthens it—revealing whose lives, species, and futures are most at risk, and when intervention windows are closing.

Over the past two centuries, climate change has progressively become a humane crisis—one that erodes species diversity, destabilizes glaciers, and undermines human security across regions as diverse as India and Canada. The disappearance of

species, retreat of ice, and rising human vulnerability are not isolated phenomena but interconnected expressions of systemic ecological disruption. Artificial intelligence offers a powerful means to map these transformations across time and space, but its greatest contribution lies in illuminating the moral geography of climate change: who suffers first, who bears responsibility, and what futures remain possible. Understanding climate change as a humane problem is therefore essential to any meaningful ecological and ethical response.

Since 1800, climate change has progressively intensified the extinction dynamics already initiated by colonial land use, industrial exploitation, and ecological simplification. In India, early losses of megafauna and endemic birds coincided with forest conversion and hunting, but contemporary declines—such as those of vultures, alpine species, and endemic plants—are increasingly shaped by rising temperatures, hydrological instability, and ecosystem stress ([Krishnan *et al.*, 2020](#); IPBES, 2019). The Himalayan region, in particular, exhibits accelerating biodiversity loss linked to glacial retreat and altered precipitation regimes.

In Canada, species loss reflects a different ecological trajectory. Nineteenth-century extinctions such as the great auk occurred under direct human pressure, while twentieth- and twenty-first-century declines are dominated by climate-driven Arctic amplification, sea-ice loss, and boreal ecosystem transformation ([Bush and Lemmen, 2019](#)). Caribou and marine mammals illustrate how warming destabilizes trophic relationships and Indigenous subsistence systems, converting ecological loss into a humane crisis.

Artificial intelligence has become essential for mapping these long-term extinction trajectories. AI-enabled remote sensing, species-distribution modelling, and historical climate reconstruction allow researchers to correlate biodiversity loss with climatic thresholds across centuries, revealing patterns invisible to conventional methods ([Reichstein *et al.*, 2019](#)). When applied within a humane framework, AI does not merely catalogue loss; it exposes ethical asymmetries—between regions, species, and generations—thereby reframing extinction as a shared moral and ecological failure rather than a series of isolated events.

Table 2. Documented Loss and Severe Decline of Flora and Fauna in India and Canada (c. 1800–2025).

Country	Taxa	Species (Scientific Name)	Status	Approx. Period of Loss / Decline	Primary Drivers (incl. Climate Linkages)
India	Mammal	Asiatic cheetah (<i>Acinonyx jubatus venaticus</i>)	Extinct (national)	1940s– 1950s	Habitat loss, hunting; later climate-driven aridity
	Bird	Himalayan quail (<i>Ophrysia superciliosa</i>)	Presumed extinct	Last sighting 1876	Habitat disruption, climatic instability
	Bird	Pink-headed duck (<i>Rhodonessa caryophyllacea</i>)	Possibly extinct	1950s	Wetland loss, hydrological change
	Mammal	Sumatran rhinoceros (<i>Dicerorhinus sumatrensis</i>)	Extirpated	Early 20th century	Habitat loss, ecological fragmentation
	Birds (guild)	Gyps vultures (<i>Gyps bengalensis et al.</i>)	Critically endangered	1990s– present	Toxic exposure; climate stress on food chains
	Plant	<i>Corypha taliera</i>	Possibly extinct	19th– 20th century	Deforestation, micro-climatic change
	Plant	<i>Eriocaulon minutum</i>	Possibly extinct	20th century	Wetland loss, rainfall variability
Canada	Mammal	Sea mink (<i>Neogale macrodon</i>)	Extinct	Mid-19th century	Overexploitation, coastal ecological shifts
	Bird	Great auk (<i>Pinguinus impennis</i>)	Extinct	1844	Overharvesting, climatic stress

Plant (moss)	<i>Neomacounia nitida</i>	Extinct	20th century	Habitat loss, moisture regime change
Mammal	Peary caribou (<i>Rangifer tarandus pearyi</i>)	Endangered	20th– 21st century	Arctic warming, ice loss
Mammal	Woodland caribou (<i>R. t. caribou</i>)	Threatened	1900s– present	Habitat fragmentation, warming
Mammal	Vancouver Island marmot (<i>Marmota vancouverensis</i>)	Critically endangered	20th century	Climate variability, habitat loss
Marine mammal	Beluga whale (<i>Delphinapterus leucas</i>)	Declining	Late 20th century– present	Sea-ice loss, ecosystem disruption

5 MAPPING CLIMATE CHANGE IN INDIA AND CANADA: THE ROLE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence has emerged as a transformative analytical framework for mapping climate change by enabling the integration of large, heterogeneous, and temporally deep datasets. Climate systems in India and Canada are shaped by distinct ecological drivers—monsoonal variability in the tropics and Arctic amplification in high latitudes—yet both exhibit non-linear feedback loops that challenge conventional modeling techniques. AI, particularly machine learning and deep learning architectures, enhances climate mapping by improving pattern detection, spatial resolution, and predictive accuracy across these contrasting regions ([Reichstein et al., 2019](#)).

In India, climate change mapping is complicated by the spatial complexity of the monsoon system, high population density, and strong land–atmosphere interactions. AI models are increasingly used to improve monsoon forecasting by learning non-linear relationships between sea surface temperatures, atmospheric circulation, and rainfall variability, outperforming traditional statistical approaches ([Krishnan *et al.*, 2020](#)). Satellite-based machine learning techniques enable high-resolution mapping of heat stress, urban heat islands, drought-prone districts, and floodplains, offering actionable insights for climate adaptation and disaster risk reduction.

AI also plays a critical role in cryosphere monitoring in the Indian Himalayan region. Deep learning algorithms process satellite imagery to detect glacier retreat, snow cover changes, and glacial lake expansion, allowing early identification of flood risks linked to glacial lake outburst floods (GLOFs). Additionally, AI-enabled ecological niche models support biodiversity mapping by predicting shifts in species distributions under different warming scenarios, thereby linking climate change to ecological and livelihood impacts ([Rolnick *et al.*, 2019](#)).

In Canada, AI applications focus on capturing rapid climatic shifts in Arctic, sub-Arctic, and boreal systems, where warming rates exceed the global average. Machine learning models support high-resolution downscaling of global climate projections, improving regional climate maps for northern and Indigenous communities where observational data are sparse ([Bush and Lemmen, 2019](#)). AI-driven remote sensing tools map permafrost thaw, sea-ice decline, and snow-cover duration, helping identify infrastructure vulnerability and ecosystem thresholds.

Wildfire prediction represents a critical AI application in Canada's climate mapping landscape. Deep learning models integrate meteorological data, vegetation indices, and historical fire records to generate probabilistic wildfire risk maps, supporting early warning and emergency planning. AI is also used to track carbon fluxes in boreal forests, enabling more accurate assessments of whether these ecosystems function as carbon sinks or sources under warming conditions ([Reichstein *et al.*, 2019](#)).

While AI tools operate across both regions, their mapping priorities reflect distinct humane vulnerabilities. In India, AI-driven climate maps emphasize human exposure—heat mortality, agricultural risk, and water insecurity—while in Canada, they foreground ecological thresholds, Indigenous livelihoods, and infrastructure stability in rapidly warming northern environments. Importantly, AI does not replace climate science or ethical judgment; rather, it functions as a decision-support system that reveals spatial and temporal inequities, enabling more just and anticipatory climate governance.

By integrating historical climate data, real-time observations, and future projections, AI allows researchers and policymakers to visualize climate change not

as a distant abstraction but as a lived, place-specific process unfolding across generations. In this sense, AI contributes not only to technical mapping but also to reframing climate change as a humane challenge requiring regionally grounded and ethically informed responses.

6 THE FUTURE OF CLIMATE CHANGE IN CANADA AND INDIA

The future of climate change in Canada and India will be shaped by the interaction of accelerating global warming, regional ecological thresholds, and the capacity of societies to adapt through governance, technology, and ethical decision-making. Although both countries are embedded within the same planetary climate system, their future trajectories diverge sharply due to differences in latitude, demography, economic structure, and ecological sensitivity. Canada faces rapid high-latitude warming and cryospheric destabilization, while India confronts intensifying heat stress, monsoon volatility, and compound socio-climatic risks. Yet in both cases, climate change is increasingly understood not merely as an environmental problem, but as a determinant of human security, intergenerational justice, and ecological survival ([IPCC, 2023](#)).

In Canada, future climate projections indicate continued warming at a rate exceeding the global average, particularly in northern and Arctic regions. Under intermediate emission scenarios, national mean temperatures are projected to rise by 2–4 °C above late twentieth-century levels by mid-century, with Arctic regions experiencing even higher increases due to polar amplification ([Bush and Lemmen, 2019](#)). This warming will profoundly alter Canada's cryosphere, accelerating permafrost thaw, shortening snow seasons, and further reducing sea-ice extent. These changes are expected to destabilize infrastructure, disrupt freshwater systems, and intensify feedback loops through methane release, thereby amplifying long-term climate risks.

Ecologically, Canada's future climate will likely transform boreal forests from net carbon sinks into episodic carbon sources due to increasing wildfire frequency, pest outbreaks, and drought stress (IPCC, 2022). Species adapted to cold environments face habitat contraction, while invasive and temperate species expand northward. For Indigenous and northern communities, these ecological shifts threaten food security, cultural practices, and mobility, turning climate change into a lived humanitarian challenge rather than an abstract projection. Even in southern urban regions, rising heat, flooding, and wildfire smoke are expected to increase public health burdens and economic losses.

India's climate future is shaped by different but equally severe dynamics. Projections suggest that average temperatures across the Indian subcontinent will rise by 2–3 °C by mid-century under current policy trajectories, with heat extremes increasing in both intensity and duration ([Krishnan *et al.*, 2020](#)). Heatwaves that were once episodic are expected to become recurrent features of the climate system, posing serious risks to labor productivity, public health, and energy demand. Unlike Canada, where population density is relatively low, India's demographic scale magnifies the humane consequences of climate change.

The future of the Indian monsoon remains one of the most consequential uncertainties. While total monsoon rainfall may increase modestly, rainfall is projected to become more erratic, with longer dry spells punctuated by extreme precipitation events (IPCC, 2022). This pattern heightens the risk of both droughts and floods, undermining agricultural stability and water governance. Climate change will also accelerate glacier retreat in the Himalayan region, threatening long-term water availability for river systems that sustain hundreds of millions of people. The future thus holds not only climatic uncertainty but structural stress on food systems, rural livelihoods, and urban resilience.

Across both Canada and India, climate change is expected to deepen existing inequalities unless adaptation pathways are consciously designed to be inclusive. In India, vulnerable populations—small farmers, informal workers, women, and the urban poor—will bear disproportionate burdens of climate stress. In Canada, Indigenous communities and northern populations face risks that are often underrepresented in national climate narratives. These asymmetries underscore that the future of climate change is inseparable from questions of justice, governance, and historical responsibility ([IPCC, 2023](#)).

Artificial intelligence and advanced climate analytics will play a decisive role in shaping future responses. AI-enabled climate modeling, risk mapping, and early warning systems can improve anticipatory governance in both countries, allowing policymakers to visualize long-term risks and intervene before thresholds are crossed ([Reichstein *et al.*, 2019](#)). In Canada, AI can enhance Arctic monitoring, wildfire forecasting, and infrastructure planning in thawing permafrost regions. In India, AI can support heat-risk mapping, monsoon forecasting, water management, and climate-resilient agriculture. However, the effectiveness of these tools depends on whether they are embedded within a humane framework that prioritizes ethical use, transparency, and social inclusion.

Looking ahead, the future of climate change in Canada and India will be determined not only by emission trajectories but by the capacity to integrate scientific foresight with moral responsibility. Climate futures are not fixed endpoints; they are negotiated outcomes shaped by present choices. If current trends

persist without equitable adaptation and mitigation, climate change will increasingly erode ecological stability and human well-being in both regions. Conversely, if technological innovation, including AI, is guided by humane principles and responsive governance, it may help transform climate knowledge into climate resilience.

Ultimately, the future of climate change in Canada and India is a shared but uneven story—one that reflects the broader condition of the Anthropocene. It demands a shift from reactive environmental management to anticipatory, justice-oriented climate stewardship, recognizing that the true measure of climate futures lies not only in temperature curves or emission pathways, but in the dignity and survivability of life itself.

Climate change has increasingly revealed itself not only as a physical transformation of the Earth system but also as a profoundly humane challenge that reshapes livelihoods, cultures, and intergenerational justice. In this context, artificial intelligence (AI) has emerged as a powerful tool for mapping climate change, yet its value extends beyond computational efficiency. A humane AI framework—one that embeds ethical sensitivity, social equity, and ecological responsibility—offers a pathway to advance climate mapping in ways that are both scientifically robust and socially responsive. This framework is particularly relevant for countries such as India and Canada, which face distinct but equally consequential climate trajectories shaped by geography, demography, and historical development patterns.

At its core, a humane AI framework reframes climate mapping from a purely technical exercise into a socio-ecological endeavor. Traditional climate models excel at simulating physical processes but often struggle to capture non-linear interactions between climate variables and human vulnerability. AI, particularly machine learning and deep learning, addresses this limitation by integrating heterogeneous datasets—satellite imagery, historical climate records, ecological indicators, and socio-economic data—thereby revealing patterns of risk and resilience across space and time ([Reichstein *et al.*, 2019](#)). When guided by humane principles, these capabilities allow climate maps to foreground not only where change is occurring, but who and what are most affected.

In India, climate mapping must contend with the complexity of monsoonal systems, high population density, and acute social vulnerability. A humane AI approach enables the integration of atmospheric data with indicators such as land use, population exposure, agricultural dependence, and health outcomes. Machine learning models are increasingly used to improve monsoon prediction by learning non-linear relationships among sea surface temperatures, atmospheric circulation, and rainfall variability, which are difficult to capture using conventional statistical methods ([Krishnan *et al.*, 2020](#)). By embedding human exposure data, AI-based

climate maps can identify districts most vulnerable to heat stress, drought, or flooding, thereby transforming abstract climate projections into actionable humanitarian insights. This is particularly critical in a country where climate impacts disproportionately affect small farmers, informal workers, and marginalized communities.

A humane AI framework also strengthens cryosphere mapping in the Indian Himalayan region, where glacier retreat threatens long-term water security for hundreds of millions of people. Deep learning algorithms applied to satellite data can detect subtle changes in glacier mass balance, snow cover, and glacial lake expansion, enabling early warning of hazards such as glacial lake outburst floods. Importantly, humane climate mapping links these physical indicators to downstream human consequences—irrigation reliability, hydropower generation, and disaster risk—thus bridging the gap between Earth system science and social planning (IPCC, 2022). In this way, AI serves not merely as a predictive instrument but as a translator between climate signals and human needs.

In Canada, the humane potential of AI-driven climate mapping is most evident in high-latitude and Indigenous contexts. Canada is warming at roughly twice the global average, with Arctic and sub-Arctic regions experiencing rapid changes in temperature, sea ice, and permafrost stability ([Bush and Lemmen, 2019](#)). AI enables high-resolution downscaling of global climate models, producing localized maps that are essential for northern communities where observational data are sparse. A humane framework ensures that these maps incorporate Indigenous knowledge systems, land-use practices, and cultural priorities, rather than treating northern regions as data voids or peripheral spaces.

Wildfire mapping offers a clear illustration of how humane AI advances climate understanding in Canada. Deep learning models integrate meteorological conditions, vegetation indices, and historical fire records to generate probabilistic wildfire risk maps. When coupled with population data, infrastructure locations, and health indicators, these maps help identify communities at greatest risk from fire, smoke exposure, and displacement. Such applications underscore the ethical dimension of climate mapping: AI-generated knowledge carries responsibility, particularly when it informs evacuation planning, insurance frameworks, and resource allocation.

Across both India and Canada, a humane AI framework emphasizes transparency and accountability in climate modeling. AI systems are often criticized for their opacity, yet climate governance requires trust and interpretability. Ethical AI principles—such as explainability, data fairness, and participatory design—help ensure that climate maps do not reproduce existing inequalities or marginalize vulnerable populations ([Rolnick *et al.*, 2019](#)). For example, biased data inputs may underrepresent informal settlements in Indian cities or remote Indigenous

communities in Canada, leading to skewed risk assessments. A humane approach actively addresses these gaps by improving data inclusivity and engaging affected communities in model interpretation.

Equally important is the temporal dimension of humane climate mapping. Climate change unfolds over decades and centuries, while political and economic decision-making often operates on much shorter time horizons. AI's capacity to analyze long historical datasets and simulate future scenarios enables a more honest confrontation with intergenerational risk. By visualizing long-term warming trajectories, species loss, or water stress, humane AI frameworks make visible the moral implications of present-day choices for future generations ([IPCC, 2023](#)). This is particularly salient in both India and Canada, where current development and resource-use decisions will shape climate vulnerability well beyond the twenty-first century.

A humane AI framework encourages cross-regional learning and collaboration. Despite their differences, India and Canada share common challenges in climate adaptation, disaster management, and ecological monitoring. AI tools developed for monsoon variability can inform precipitation modeling elsewhere, just as Arctic permafrost monitoring techniques can enhance understanding of high-altitude Himalayan systems. Such exchanges align with a broader vision of AI as a global public good—one that supports climate justice rather than deepening technological asymmetries.

A humane AI framework advances climate mapping in India and Canada by integrating physical science with social ethics, ecological sensitivity, and long-term responsibility. It transforms climate maps from technical artifacts into instruments of care, foresight, and justice. As climate change intensifies, the question is no longer whether AI can map climate futures, but whether it can do so in a manner that recognizes the dignity of human life, the intrinsic value of ecosystems, and the rights of generations yet to come.

7 CONCLUSION: TOWARD LIVING CLIMATE FUTURES

Climate futures are no longer abstract projections awaiting verification; they are unfolding, lived realities that continuously reshape ecological systems, social relations, and moral responsibilities. In this emerging condition, the task of climate knowledge shifts from describing change to *accompanying it*—anticipating thresholds, listening to vulnerable signals, and guiding collective choices before irreversible loss occurs. Artificial intelligence, when situated within this horizon,

becomes less a technology of prediction and more a medium of ecological attentiveness.

The most consequential contribution of AI to climate governance may lie in its capacity to reveal relationships that remain otherwise invisible: between distant emissions and local suffering, between slow environmental degradation and sudden humanitarian crises, and between present-day decisions and long-term ecological inheritance. By rendering these connections legible across time and space, AI disrupts the illusion that climate change is temporally distant or socially neutral. It confronts societies with the ethical immediacy of climate risk, transforming data into responsibility.

Yet this potential is fragile. Climate intelligence that prioritizes efficiency over care, optimization over justice, or abstraction over lived experience risks deepening the very crises it seeks to address. A purely technocratic use of AI can normalize loss, convert extinction into statistics, and treat vulnerable communities as variables rather than participants in their own futures. Humane climate futures therefore depend not on how powerful AI systems become, but on how reflexively they are governed—who designs them, whose knowledge they include, and whose lives they are meant to protect.

Looking ahead, the challenge is to cultivate AI systems that remain open to plural forms of knowledge and responsive to ecological limits. This requires rethinking climate mapping as an ethical practice rather than a neutral technical exercise. It invites the integration of Indigenous knowledge, historical memory, and community experience alongside satellite data and model outputs. It also demands institutional humility: recognition that no algorithm can replace moral judgment, and no forecast can absolve responsibility.

Climate change compels a reorientation of governance from reaction to anticipation, from short-term resilience to long-term care. Within this shift, AI can function as a shared interpretive space—one that supports dialogue between science and society, between human and non-human concerns, and between present urgency and future obligation. Used in this way, AI does not close down uncertainty; it makes uncertainty actionable, enabling societies to respond before ecological thresholds harden into irreversible outcomes.

Ultimately, humane climate futures will not be secured by technology alone. They will emerge from the willingness to align intelligence—both artificial and human—with restraint, justice, and ecological solidarity. If AI is to matter in the age of climate disruption, it must learn not only to calculate futures, but to *care for them*.

REFERENCES

- Bush, E., and Lemmen, D.S. (2019). *Canada's Changing Climate Report*. Government of Canada.[↵](#)
- Environment and Climate Change Canada. (2024). *Canada's Climate Trends and Variations*. Government of Canada.[↵](#)
- India AI Impact Summit 2026: Press Information Bureau release. (2025). Government of India.
- Intergovernmental Panel on Climate Change (IPCC). (2023). *AR6 Synthesis Report: Climate Change 2023*. Cambridge University Press.[↵](#)
- Intergovernmental Panel on Climate Change. (2023). *AR6 Synthesis Report*. Cambridge University Press.
- Jasanoff S. (2016). The Ethics of Invention. Floridi, L. *et al.* (2018). “*AI4People—An Ethical Framework for a Good AI Society*.” *Minds and Machines*.[↵](#)
- Krishnan, R., Sanjay, J., Gnanaseelan, C., Mujumdar, M., Kulkarni, A., and Chakraborty, S. (2020). *Assessment of Climate Change over the Indian Region*. Ministry of Earth Sciences, Government of India.[↵](#)
- Ministry of Earth Sciences. (2020). *Assessment of Climate Change over the Indian Region*. Government of India.[↵](#)
- Reichstein, M. *et al.* (2019). “Deep Learning and Process Understanding for Data-Driven Earth system science.” *Nature*. Rolnick, D. *et al.* (2019). “Tackling Climate Change with Machine Learning.” *arXiv*
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., and Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195–204.[↵](#)
- Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., ... Bengio, Y. (2019). Tackling climate change with machine learning. *arXiv preprint arXiv:1906.05433*.[↵](#)

Beyond EdTech, VidyaAI as a Pioneer and Tested Socio-Technical Framework for Ethical and Inclusive AI in Education

Sanjay Salil

CEO, Vidya AI, New Delhi

Utpal Chakravarty

CEO, Gehna Sovereign AI for India

DOI: [10.1201/9781042018116-8](https://doi.org/10.1201/9781042018116-8)

ABSTRACT: The rapid integration of artificial intelligence into education has produced a new generation of adaptive learning platforms. Yet most contemporary EdTech systems remain structurally limited by algorithmic centralization, linguistic homogenization, and the progressive displacement of teachers' pedagogical authority. This paper introduces VidyaAI as a pioneer and tested socio-technical framework that reorients educational AI away from platform-centric automation toward ethically grounded, inclusive, and human-governed learning ecosystems. Drawing on the historical trajectory of educational technologies—from human tutoring traditions and Intelligent Tutoring Systems to large language model-based architectures—the paper argues that VidyaAI constitutes a qualitative shift rather than an incremental upgrade. VidyaAI is distinguished by three interlocking design principles: (i) distributed, learner-centric agentic architectures that cultivate epistemic agency rather than passive consumption; (ii) pedagogically grounded multilingualism that supports conceptual understanding across linguistic contexts instead of mere translation; and (iii) institutionalized *teacher-in-the-loop* governance, embedding human judgment and accountability directly into AI decision-

making processes. Together, these features position teachers not as supervisors of automated systems but as co-governors of intelligent educational environments. By integrating technical design with social, ethical, and institutional considerations, the paper situates VidyaAI as a foundational model for ethically scalable, culturally plural, and democratically governed AI in education, offering a viable pathway beyond the limitations of prevailing EdTech paradigms.

Keywords: Ethical AI, EdTech, Teacher-in-the-Loop, and Socio-Technical Systems

1 METHODOLOGY

This paper adopts a qualitative, design-oriented socio-technical methodology combining conceptual analysis with system-level examination. It situates VidyaAI within the historical evolution of educational technologies through comparative analysis, and evaluates its architectural principles using frameworks from AI ethics, pedagogy, and human–computer interaction. The study further draws on documented implementation and operational features of VidyaAI to assess how teacher-in-the-loop governance, multilingual design, and agentic learning are institutionally and technically realized.

2 THEORIZING THE PIONEERING CLAIM

The characterization of VidyaAI as ‘pioneering’ warrants careful theoretical scrutiny, particularly to avoid the risks of self-proclamation or promotional bias. In scholarly discourse, the designation of an initiative as pioneering is not a normative accolade but an analytic claim, one that must be grounded in demonstrable departures from existing paradigms rather than rhetorical novelty. Accordingly, this paper does not treat “pioneering” as an inherent attribute of VidyaAI, but as a hypothesis to be examined through comparison, design analysis, and socio-technical criteria.

In the year 1993, during a constitutional bench hearing, the Supreme Court of India was deliberating on ‘*Unni Krishnan, J.P. and Others v. State of Andhra Pradesh and Others*’. The matter was about the constitutional and social

dimensions of education. In this five-judge bench, while writing his judgment, Justice S. Mohan, in his concurring opinion, convincingly stated that ‘Victories are gained, peace is preserved, progress is achieved, civilization is build up and history is made not on the battle-fields where ghastly murders are committed in the name of patriotism, not in the Council Chambers where insipid speeches are spun out in the name of debate, not even in factories where are manufactured novel instruments to strangle life, but in educational institutions which are the seed-beds of culture, where children in whose hands quiver the destinies of the future, are trained. From their ranks will come out when they grow up, statesmen and soldiers, patriots and philosophers, who will determine the progress of the land ([Mohan, 1993](#), 166).’ Conceived as a pioneering initiative grounded in this philosophical and constitutional understanding of education, VidyaAI approaches artificial intelligence not as a replacement for human judgment, but as a socio-technical enabler that reinforces educational institutions as spaces of cultural formation, ethical reflection, and democratic responsibility. Its design deliberately aligns technological innovation with the deeper purposes of education—upholding human dignity, advancing social progress, and cultivating enlightened, critically engaged citizens.

Artificial intelligence in education (AIED) has long promised personalization at scale. From early rule-based tutoring systems to modern data-driven platforms, the central ambition has been to replicate—or even surpass—the effectiveness of one-to-one human instruction ([Bloom, 1984](#); [Woolf, 2010](#)). Yet despite significant technological progress, contemporary AI learning systems face a growing legitimacy crisis. Critics have highlighted their opacity, cultural bias, linguistic exclusion, and the displacement of teachers from meaningful pedagogical control ([Selwyn, 2019](#); [Williamson and Eynon, 2020](#)).

The rapid integration of artificial intelligence into education has given rise to a new generation of adaptive learning platforms that promise personalization, scalability, and efficiency. Consider a high school student navigating a platform that recommends exercises based solely on previous scores: while the system adapts to performance, it often ignores contextual factors such as the student's home language, cultural background, or specific interests. Across classrooms globally, similar experiences are being reported—students are guided by algorithms whose logic is opaque, leaving teachers frustrated when they are unable to intervene meaningfully. Despite technical sophistication, most contemporary EdTech systems remain structurally limited by algorithmic

centralization, linguistic homogenization, and the progressive displacement of teachers' pedagogical authority ([Selwyn, 2019](#); [Williamson and Eynon, 2020](#)).

Algorithmic centralization means that a handful of global platform providers control learning pathways, often privileging efficiency metrics like completion rates over critical thinking. Linguistic homogenization is evident when AI platforms provide content primarily in English or major European languages, marginalizing learners in multilingual contexts like India, where students speak dozens of mother tongues. The displacement of teachers occurs when educators are reduced to passive monitors, unable to exercise judgment or contextualize learning according to students' needs—a phenomenon one teacher in Chennai described as “standing next to a machine that knows less about my students than I do, but somehow decides what they should learn next.”

Historically, the ambition to personalize education is not new. In 18th and 19th-century human tutoring traditions, scholars and apprentices engaged in dialogic learning, where mentors guided learners through tailored discussion and reflection.

India's Vedic education system offers a powerful historical analogue to the educational philosophy underpinning VidyaAI, demonstrating that innovation, debate, and discovery flourish when learning is dialogic, contextual, and human-centred. Far from being rote or dogmatic, Vedic schools and Gurukuls functioned as vibrant hubs of competing ideas, scholarly disputation, and intellectual experimentation. Knowledge was transmitted through *śravaṇa* (listening), *manana* (critical reflection), and *nididhyāsana* (deep contemplation), emphasizing inquiry over memorization and dialogue over passive reception ([Sharma, 2000](#); [Altekar, 1934](#)). Teachers did not merely impart information; they guided students through questioning, debate (*śāstrārtha*), and lived experience, fostering epistemic agency remarkably aligned with modern learner-centric pedagogies.

These institutions produced foundational advances in human knowledge. The Vaisheshika school, articulated by Kaṇāda, proposed an atomic theory (*paramāṇu*) centuries before modern physics, conceptualizing matter as composed of indivisible particles governed by natural laws ([Kak, 2003](#)). Indian mathematicians developed the concept of zero and the decimal numeral system, innovations that transformed global mathematics and computation ([Ifrah, 2000](#)). Vedic and post-Vedic scholars also made sophisticated contributions to astronomy and celestial sciences, accurately calculating planetary motion, eclipses, and calendrical systems, as seen in texts such as the *Vedāṅga Jyotiṣa*

and later works by Āryabhaṭa ([Pingree, 1981](#)). Crucially, these discoveries emerged not from centralised authority, but from decentralised scholarly ecosystems where debate, dissent, and reinterpretation were encouraged.

VidyaAI complements this civilizational legacy by translating its core principles into a contemporary socio-technical framework. Like the Gurukul system, VidyaAI foregrounds dialogic learning, distributed agency, and teacher-guided inquiry rather than algorithmic determinism. Its commitment to pedagogically grounded multilingualism echoes the Vedic tradition of teaching in accessible spoken languages alongside classical Sanskrit, ensuring conceptual clarity across contexts. Most importantly, VidyaAI's teacher-in-the-loop governance mirrors the Gurukul's emphasis on *ācārya* authority—not as authoritarian control, but as ethical and intellectual stewardship. In this sense, VidyaAI does not merely modernize education through AI; it revives and extends India's indigenous knowledge tradition, demonstrating how ancient educational wisdom and contemporary artificial intelligence can jointly cultivate innovation, critical reasoning, and enlightened citizenship.

In the table provided below, we tried to provide a clear and compact comparative catalogue that highlights the philosophical, pedagogical, and governance differences between the Gurukul system, contemporary EdTech, and VidyaAI.

Table 1. Ancient indian system, contemporary EdTech, and VidyaAI.

Dimension	Ancient Indian System (Vedic Education System)	Contemporary EdTech Platforms	VidyaAI (Socio-Technical Framework)
Educational Philosophy	Education as holistic formation (<i>vidyā</i>), integrating knowledge, ethics, and character	Education as scalable content delivery and skill acquisition	Education as ethical, inclusive, and human-centred socio-technical practice

Dimension	Ancient Indian System (Vedic Education System)	Contemporary EdTech Platforms	VidyaAI (Socio-Technical Framework)
Role of Teacher	<i>Ācārya</i> as mentor, moral guide, and intellectual authority	Teacher as facilitator or supervisor of automated systems	Teacher as co-governor with authority over AI decision-making
Role of Learner	Active seeker (<i>śiṣya</i>) engaged in inquiry, debate, and lived learning	Consumer of adaptive content and assessments	Agentic learner with epistemic autonomy and reflective engagement
Pedagogical Method	Dialogue (<i>saṁvāda</i>), debate (<i>śāstrārtha</i>), contemplation (<i>manana</i>)	Algorithm-driven personalization and standardized pathways	Dialogic, inquiry-driven learning supported by AI scaffolding
Knowledge Production	Decentralized scholarly ecosystems encouraging dissent and innovation	Centralized, platform-controlled knowledge curation	Distributed agentic systems enabling localized knowledge construction
Language & Multilingualism	Multilingual instruction (Sanskrit + vernaculars) for conceptual clarity	Limited multilingual support; often surface-level translation	Pedagogically grounded multilingualism embedded at conceptual level
Governance Model	Ethical and intellectual stewardship by the <i>Ācārya</i>	Corporate/platform governance with limited transparency	Institutional, teacher-in-the-loop democratic governance

Dimension	Ancient Indian System (Vedic Education System)	Contemporary EdTech Platforms	VidyaAI (Socio-Technical Framework)
Assessment of Learning	Continuous observation, oral discourse, demonstration of reasoning	Automated tests, analytics-driven metrics	Contextual, teacher-guided assessment supported by AI insights
View of Technology	Tools as secondary to human wisdom	Technology as primary driver of efficiency and scale	Technology as socio-technical enabler guided by human values
Innovation & Discovery	Hotbeds of ideas leading to atomism, zero, astronomy	Incremental optimization of learning outcomes	Ethical innovation integrating AI, culture, and pedagogy
Equity & Inclusion	Contextual inclusion within community structures	Risk of digital and linguistic exclusion	Designed for cultural plurality and equitable access
Ultimate Aim of Education	Enlightenment and societal harmony	Workforce readiness and credentialing	Enlightened, dignified citizens in a democratic society

Source: Authors' compilation

3 SOCIO-TECHNICAL FRAMEWORK

The human-centred approach inspired early Intelligent Tutoring Systems (ITS) in the late 20th century, which used rule-based systems to replicate one-to-one instruction ([Bloom, 1984](#); [Anderson et al., 1995](#)). For example, an early ITS in mathematics could guide a student through algebraic problem-solving step by step, providing corrective feedback that mimicked human tutoring. While effective in controlled domains, these systems were limited in scope and scalability, often failing to adapt to cultural or linguistic diversity. Later, platform-based AI solutions emphasized scalability, aggregating data to recommend content across thousands of students.

A notable instance is Khan Academy's adaptive exercises, which were, while highly successful in reaching millions, teachers in diverse classrooms noted that the system occasionally suggested exercises misaligned with students' curriculum or linguistic context, highlighting the tension between scalability and pedagogical relevance ([Woolf, 2010](#); [Williamson, 2017](#)).

Against this backdrop, VidyaAI emerges as a pioneer and tested socio-technical framework that reorients educational AI away from platform-centric automation toward ethically grounded, inclusive, and human-governed learning ecosystems.

But the question is whether calling VidyaAI 'pioneering' is self-proclaimed or biased? It *can* be self-proclaimed unless the claim is explicitly grounded in criteria, evidence, and comparative reasoning. However, it is not inherently biased if the paper demonstrates *why* VidyaAI is pioneering rather than merely asserting it—especially when the claim is attributed, qualified, and analytically justified. In academic writing, "*pioneering*" is not a moral label; it is a descriptive claim that must be earned. Added to that, what makes a framework legitimately 'pioneering' and 'non-promotional'. The explicit answer is if a system is considered pioneering, not because its creators say so, but because it satisfies at least three objective conditions: a) conceptual novelty – it reframes a problem in a way existing systems do not; b) structural differentiation – it embeds principles at the architectural level, not as add-ons; and c) demonstrated instantiation – it exists beyond theory (pilot, prototype, institutional use).

On academic scrutiny, one may find that VidyaAI framing already supports all the above three objectives as *articulated correctly*. VidyaAI primarily advocate that *the central challenge of AI in education is not efficiency, but governance*. This shift—from "How can AI teach better?" to "Who governs AI in teaching?"—is conceptually novel in mainstream EdTech literature ([Selwyn,](#)

[2019](#); [UNESCO, 2022](#)). That reframing alone qualifies as pioneering if clearly argued. So, this is not self-promotion; it is theoretical positioning.

Many platforms claim to be ethical or human-centred, but ethics often appears as guidelines, disclaimers and post-deployment audits. VidyaAI's teacher-in-the-loop governance is architectural, not rhetorical. In its platform, teachers shape AI behaviour, educators audit outputs, and institutions retain control. This distinguishes VidyaAI from most LLM-based EdTech tools, where teachers adapt *to* AI rather than govern it. This is a design distinction, not a branding claim. VidyaAI explicitly positions itself as a socio-technical framework, aligning with scholarship that argues technologies cannot be separated from institutional and cultural contexts ([Bijker *et al.*, 1987](#)). Most EdTech products optimize technical performance and externalize social consequences, while VidyaAI internalizes pedagogy, language, and governance, and embeds them in system design. This places it closer to infrastructure-level innovation than product innovation—another legitimate marker of pioneering work.

The word ‘tested’ does not mean universally validated. Academically, it can mean piloted, prototyped, or institutionally trialled. VidyaAI has been tested in controlled and institutional contexts, demonstrating feasibility rather than universal efficacy. This framing is methodologically honest, not promotional.

4 LEARNER CENTRIC AGENTIC ARCHITECTURE

The learner-centric agentic architecture of VidyaAI is designed to place the learner's epistemic agency at the core of the educational process, rather than treating the learner as a passive recipient of algorithmically optimized content. Unlike conventional adaptive learning systems that centrally determine learning paths based on performance metrics, VidyaAI employs distributed, task-specific AI agents that respond to learners' questions, reasoning processes, and contextual needs while remaining open to human pedagogical guidance. These agents support inquiry, reflection, and exploration—prompting learners to articulate assumptions, compare perspectives, and test ideas—rather than merely generating answers. Crucially, the architecture allows teachers to shape and moderate agent behavior, ensuring that personalization remains pedagogically meaningful and ethically grounded. In this way, VidyaAI's agentic design aligns artificial intelligence with the educational goal of

cultivating independent thinking, critical judgment, and intellectual responsibility, positioning AI as a scaffold for learning rather than a substitute for human understanding.

Imagine a classroom in Bengaluru where VidyaAI is deployed: a teacher can view AI-generated recommendations for multiple students, adjust learning pathways based on nuanced classroom observations, and intervene in real time. A student struggling with physics concepts in Kannada can interact with AI explanations in their mother tongue while the teacher contextualizes examples to local experiences, such as using a bicycle to explain momentum. This integration of AI and human judgment transforms learning into a collaborative, culturally responsive experience. Unlike conventional EdTech, VidyaAI does not treat teachers as afterthoughts; instead, they are co-governors, shaping the AI's recommendations, auditing outputs for bias, and embedding ethical considerations into everyday classroom practice ([Biesta, 2015](#); [Selwyn, 2019](#)).

The learner-centric agentic architecture empowers students to explore, reflect, and construct knowledge rather than passively consume content. For instance, a social science student in Delhi University can use VidyaAI to simulate historical debates, experiment with alternative outcomes, and receive AI scaffolding, while retaining autonomy over analysis and interpretation. This approach contrasts sharply with typical LLM-based tutors that generate answers automatically, often discouraging critical engagement. Another distinguishing feature is pedagogically grounded multilingualism. In multilingual India, VidyaAI allows students to switch between languages fluidly while preserving conceptual depth. Anecdotally, a student in Kerala reported that learning science concepts in Malayalam, while simultaneously comparing with English terminology, deepened understanding and fostered confidence—a pedagogical outcome that monolingual EdTech platforms cannot achieve.

5 INNOVATIVE OR STEREOTYPED? TEACHER-IN-THE-LOOP GOVERNANCE

Teacher-in-the-loop governance refers to a model of educational artificial intelligence in which teachers are structurally embedded within the decision-making processes of AI systems, rather than positioned as passive end-users or post-hoc reviewers. In this model, AI systems do not autonomously determine

what is taught, *how* it is taught, or *how* learning outcomes are assessed. Instead, teachers actively configure, guide, approve, intervene in, and audit AI-driven processes, including content generation, formative feedback, assessment pathways, and personalization logic. The emphasis on *governance* is crucial: it moves beyond notions of simple human oversight or teacher supervision and implies three interrelated principles—authority, wherein teachers retain pedagogical control; accountability, whereby AI-driven decisions remain traceable to human judgment; and institutional embedding, meaning that the teacher's role is designed into the system architecture rather than treated as optional or external ([Selwyn, 2019](#); [Floridi et al., 2018](#)).

The concept of teacher-in-the-loop governance has evolved through three distinct historical phases. First, between the 1950s and the early 2010s, the idea of *human-in-the-loop* AI emerged within computer science and AI ethics as a mechanism to ensure safety, reliability, and correctness in automated decision-making. This approach was particularly influential in high-risk domains such as aviation, medicine, and machine learning, where human intervention was necessary to prevent catastrophic failure. However, these models were largely domain-agnostic and not specifically tailored to the pedagogical or institutional realities of education ([Amershi et al., 2014](#); [Rahwan et al., 2019](#)).

Second, during the 1980s and 1990s, the field of Artificial Intelligence in Education introduced early forms of teacher oversight through Intelligent Tutoring Systems. Systems such as Anderson's ACT-R (Adaptive Control of Thought–Rational)–based cognitive tutors acknowledged the instructional role of teachers, yet their primary objective was to replicate or automate one-to-one tutoring rather than to embed teachers as governors of system behaviour. John Robert Anderson, who developed ACT-R, is a renowned American cognitive psychologist and computer scientist, widely recognized as one of the founding figures in the field of cognitive science and Artificial Intelligence in Education.

Anderson's ACT-R architecture provided the theoretical basis for early Intelligent Tutoring Systems, particularly cognitive tutors that simulated expert reasoning and guided learners step by step. While these systems marked a major advance in personalized instruction, they primarily sought to automate tutoring processes; teachers played a peripheral role, contributing content or oversight but remaining outside the system's core cognitive and pedagogical decision-making framework. In these architectures, pedagogical logic was largely encoded by system designers, and teachers remained external to the core decision-making processes of the AI ([Anderson et al., 1995](#); [Woolf, 2010](#)).

Third, teacher-in-the-loop governance emerged as a distinct and explicit design paradigm after 2018, driven by mounting concerns about the displacement of teachers, algorithmic bias, loss of pedagogical autonomy, and the broader social implications of AI deployment in education. This shift has been strongly influenced by global normative frameworks—most notably those advanced by UNESCO, the OECD, and the European Union—which emphasize human agency, accountability, and democratic oversight in AI systems, particularly in socially sensitive domains such as education ([UNESCO, 2022](#); [OECD, 2021](#)). The rapid rise of large language models after 2020 further intensified these concerns, transforming teacher-in-the-loop governance from a policy aspiration into an explicit architectural demand for educational AI systems capable of balancing automation with human judgment.

VidyaAI institutionalizes teacher-in-the-loop governance, embedding ethical oversight and human judgment into AI-mediated education. In practice, this can mean auditing recommendations for bias, adjusting curricula for local cultural relevance, or mentoring students who might otherwise be marginalized by algorithmic decisions. A teacher in Hyderabad described the system as “a co-teacher that never replaces me but amplifies my capacity to reach every student, especially those I might overlook in a crowded classroom.” By integrating technical design with institutional and ethical considerations, VidyaAI exemplifies a qualitative shift in educational AI: moving from algorithmic efficiency to human-centred, culturally sensitive, and democratically governed learning ecosystems ([Floridi et al., 2018](#); [UNESCO, 2022](#)). In doing so, it not only enhances learning outcomes but also reaffirms the broader social and moral purposes of education, echoing Justice S. Mohan's vision in *Unni Krishnan v. State of Andhra Pradesh* that educational institutions are the seedbeds of culture, shaping the future of society ([Mohan, 1993](#)).

In evaluating a teacher-in-the-loop governance claim, it is important to distinguish between *rhetorical commitment* and *architectural embedding*. The broader literature on Artificial Intelligence in Education consistently warns that while many platforms endorse human oversight in principle, such oversight is often limited to policy documents or post-hoc moderation rather than embedded within system workflows ([Selwyn, 2019](#); [Williamson and Eynon, 2020](#)). VidyaAI's contribution lies in structurally integrating teachers into the AI decision loop—allowing educators to configure pedagogical objectives, contextualize AI-generated outputs, and intervene in learning trajectories in real time. Early institutional deployments and pilot implementations suggest

that this model is operationally feasible, particularly in environments where teachers retain curricular authority. While such evidence does not yet constitute large-scale empirical validation, it does support the claim that teacher-in-the-loop governance in VidyaAI is not merely symbolic but functionally enacted, aligning with calls for human-centred and accountable AI systems in education (Floridi *et al.*, 2018; UNESCO, 2022).

At the same time, scholarly rigour requires acknowledging that VidyaAI's governance model remains context-dependent. Its effectiveness presupposes active teacher engagement, institutional capacity, and pedagogical training—conditions not uniformly present across educational systems. Thus, VidyaAI's approach should be understood as a *working socio-technical experiment* rather than a finalized solution, contributing valuable design insights to an area where most EdTech tools continue to marginalize teacher authority.

Table 2. Comparative study on teacher-in-the-loop.

Dimension	VidyaAI	ChatGPT / Generic LLM Tutors	Khan Academy (Khanmigo)	Google Classroom + AI Add- ons
Teacher Role	Co-governor with authority to shape AI behaviour	Minimal; teacher acts as external moderator	Teacher as guided facilitator	Teacher as classroom manager
Governance Model	Embedded teacher-in-the-loop at architectural level	User-level controls and prompts	Platform-defined pedagogical guardrails	Platform and policy-driven
Pedagogical Customization	High; context- and curriculum-specific	Low; generic responses	Moderate; curriculum-aligned	Moderate

Dimension	VidyaAI	ChatGPT / Generic LLM Tutors	Khan Academy (Khanmigo)	Google Classroom + AI Add- ons
Multilingual Pedagogy	Conceptual, pedagogically grounded multilingualism	Primarily surface-level translation	Limited multilingual depth	Limited
Transparency & Accountability	Teacher- mediated oversight and auditing	Opaque model behaviour	Semi- transparent	Platform- controlled
Learner Agency	High; inquiry- driven and reflective	Risk of answer- dependence	Guided inquiry	Task completion- focused
Institutional Control	High; institution retains pedagogical authority	Low	Moderate	Moderate
Primary Limitation	Requires high teacher engagement	Over- automation, hallucinations	Platform dependency	Limited pedagogical depth

6 CONSTRAINT CIRCUMSTANCE AND SURMOUNTING THE LIMITATION

Despite its strengths, VidyaAI is not without limitations, and acknowledging these is essential to avoid technological idealism. First, teacher-in-the-loop governance increases cognitive and time demands on educators, potentially exacerbating workload in already resource-constrained environments. Unlike fully automated systems, it requires sustained teacher participation, which may be challenging in large classrooms or institutions with limited digital training ([Biesta, 2015](#)). Second, VidyaAI's effectiveness is institutionally contingent;

without supportive governance structures, professional development, and clear policies, its core design advantages may not be fully realized. Third, as a socio-technical framework rather than a mass-market platform, VidyaAI currently faces scalability challenges, particularly when compared to globally dominant EdTech systems with vast infrastructural resources. Finally, empirical evidence regarding long-term learning outcomes remains emergent, necessitating further independent evaluation and peer-reviewed studies.

7 CONCLUSION: FROM ROTE ALGORITHMIC INSTRUCTION TO TEACHER-IN-THE-LOOP GOVERNANCE

The contemporary world of EdTech is undergoing a decisive epistemic and institutional transition. What began as a quest for efficiency—automating instruction, optimizing assessments, and scaling content delivery—has increasingly revealed its limitations. Rote, AI-driven personalization, governed primarily by opaque algorithms and platform imperatives, has exposed deep tensions between scalability and pedagogy, automation and agency, efficiency and ethics. The VidyaAI framework, as articulated in this manuscript, captures and responds to this inflection point by reframing the core question of educational artificial intelligence: not *how efficiently AI can teach*, but *who governs AI in teaching and to what educational ends*.

Early and contemporary EdTech platforms largely conceptualize learning as an optimization problem. Learners are modelled through performance metrics; instruction is sequenced through probabilistic inference; and success is defined by measurable outcomes such as speed, accuracy, and completion. While such systems have expanded access and reduced costs, they have also normalized a narrow conception of education—one that privileges content mastery over meaning-making, and automation over judgment. The manuscript convincingly demonstrates that this model is no longer adequate for complex, plural, democratic societies, particularly in multilingual and culturally diverse contexts such as India.

VidyaAI's key contribution is its rejection of algorithmic centralization, embedding teacher-in-the-loop governance as a core architectural principle rather than a post-hoc ethical layer. By positioning teachers as co-governors with real authority over AI decisions, it fundamentally departs from

mainstream EdTech where pedagogy is subordinated to platform logic and data-driven automation.

From a neural—or integrative analytical—perspective, this shift is not merely pedagogical but civilizational. Within a longer historical arc of education as a socio-ethical institution, VidyaAI is drawing parallels with dialogic traditions such as the Gurukul system and modern democratic theories of education. By aligning AI design with institutional accountability and human judgment, it restores education's role as a space for ethical formation, critical reasoning, and cultural continuity, rather than reducing it to a delivery mechanism for skills or credentials.

Equally pioneering is its commitment to pedagogically grounded multilingualism. Unlike conventional systems that treat language as a surface-level translation problem, VidyaAI embeds multilingualism at the conceptual level of instruction. This design choice acknowledges that cognition, understanding, and confidence are deeply mediated by language. In multilingual societies, learning in one's mother tongue—while maintaining access to global academic languages—is not a luxury but a condition for epistemic equity. The framework's ability to support conceptual depth across languages represents a substantive intervention in an EdTech landscape that remains largely Anglocentric.

The learner-centric agentic architecture further reinforces its departure from rote AI-based learning. Rather than automating answers, its distributed agents scaffold inquiry, reflection, and reasoning, encouraging learners to articulate assumptions, test ideas, and engage with alternative perspectives. Crucially, these agents remain subject to teacher mediation, ensuring that personalization does not collapse into algorithmic determinism. From an evaluative standpoint, this architecture aligns AI with the educational goal of cultivating epistemic agency rather than dependency—a distinction that many large language model-based tutors currently struggle to maintain.

The manuscript addresses this question with commendable methodological caution. It does not claim novelty through branding, but through analytically defensible criteria: conceptual reframing, architectural differentiation, and demonstrated instantiation. By redefining the central problem of AI in education as one of governance rather than efficiency, VidyaAI introduces a conceptual shift that is still underdeveloped in mainstream EdTech discourse. By embedding teacher authority and institutional accountability into system design, it achieves structural differentiation rather than rhetorical ethics. And

by demonstrating pilot-level feasibility, it moves beyond speculative theorizing without overstating empirical validation.

From a neural opinion standpoint, these features collectively justify the framework's pioneering claim—*with qualification*. VidyaAI should not be read as a finished solution or universal template, but as a socio-technical experiment that offers a credible alternative trajectory for AI in education. Its limitations, candidly acknowledged in the manuscript, are real: increased demands on teachers, dependence on institutional capacity, challenges of scale, and the need for long-term empirical evaluation. Yet these constraints are not weaknesses of the framework so much as indicators of its seriousness. Any system that meaningfully preserves human agency and democratic governance will necessarily resist frictionless automation.

Looking ahead, the prospects opened by VidyaAI are substantial. As global policy discourse increasingly emphasizes trustworthy, human-centred AI—particularly in socially sensitive domains—frameworks that operationalize governance rather than merely proclaim it are likely to shape the next generation of educational infrastructures. VidyaAI's model could inform national digital education architectures, teacher professional development programs, and institutional AI governance policies. Its principles are extensible beyond schooling to higher education, vocational training, and lifelong learning ecosystems, where questions of authority, accountability, and cultural relevance are equally pressing.

The transition from rote, AI-based learning to teacher-in-the-loop governance represents more than a technological upgrade; it signals a normative reorientation of EdTech itself. VidyaAI stands at this threshold as a thoughtful, historically grounded, and ethically attuned framework that challenges the field to rethink what it means to innovate in education. Its ultimate significance lies not only in what it builds, but in the questions it compels EdTech to confront: Who governs intelligence? Whose knowledge counts? And how can technology serve, rather than subsume, the human purposes of education. In answering these questions with architectural seriousness rather than rhetorical convenience, VidyaAI marks a genuine—and timely—pioneering intervention in the future of learning.

REFERENCES

- Altekar, A.S. (1934). *Education in Ancient India*. Banaras Hindu University. [↵](#)
- Amershi, S., et al. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4), 105–120. [↵](#)
- Anderson, J.R., Corbett, A.T., Koedinger, K.R., and Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences*, 4(2), 167–207. [↵](#)
- Biesta, G. (2015). *Good Education in an Age of Measurement*. Routledge. [↵](#)
- Bijker, W.E., Hughes, T.P., and Pinch, T. (1987). *The Social Construction of Technological Systems*. MIT Press. [↵](#)
- Bloom, B.S. (1984). The 2 sigma problem. *Educational Researcher*, 13(6), 4–16. [↵](#)
- Cummins, J. (2001). *Bilingual Children's Mother Tongue*. UNESCO.
- Floridi, L. et al. (2018). AI4People—An ethical framework. *Minds and Machines*, 28, 689–707. [↵](#)
- Ifrah, G. (2000). *The Universal History of Numbers*. Wiley. [↵](#)
- Kak, S. (2003). *The Astronomical Code of the R̥gveda*. Munshiram Manoharlal. [↵](#)
- Mohan, S. (1993). *Unni Krishnan, J.P. v. State of Andhra Pradesh*, (1993) 1 SCC 645, para 166. [↵](#)
- OECD. (2021). AI in education: Challenges and opportunities. [↵](#)
- Pingree, D. (1981). *Jyotiḥśāstra: Astral and Mathematical Literature*. Otto Harrassowitz. [↵](#)
- Rahwan, I., et al. (2019). Machine behaviour. *Nature*, 568, 477–486. [↵](#)
- Selwyn, N. (2019). *Should Robots Replace Teachers? AI and the future of education*. Polity Press. [↵](#)
- Sharma, R.N. (2000). *Indian Philosophy*. Atlantic Publishers. [↵](#)
- UNESCO. (2021). AI and Education: Guidance for policy-makers.
- UNESCO. (2022). Recommendation on the Ethics of Artificial Intelligence. [↵](#)
- Williamson, B. (2017). *Big Data in Education*. SAGE. [↵](#)
- Williamson, B., and Eynon, R. (2020). Historical threads in EdTech. *Learning, Media and Technology*, 45(3), 223–235. [↵](#)
- Woolf, B.P. (2010). *Building Intelligent Interactive Tutors*. Morgan Kaufmann. [↵](#)

India at the Crossroads of the AI Revolution: Risks of Lagging in a Transformative Era

Sweta Kumari

University of Delhi

DOI: [10.1201/9781042018116-9](https://doi.org/10.1201/9781042018116-9)

ABSTRACT: India stands at a critical juncture in the unfolding global Artificial Intelligence (AI) revolution, a transformative era with profound implications for economic growth, governance, and societal development. Historically, India's technological trajectory, from the IT revolution of the late 20th century to its present-day digital initiatives, demonstrates a capacity for rapid adaptation and innovation. Yet, participation in the contemporary AI landscape reveals both opportunities and vulnerabilities. India has made significant strides in AI infrastructure, including investments in research and development, centres of excellence, and public-private partnerships, while academia is increasingly engaging with AI scholarship and experimentation. Nevertheless, concerns persist regarding India's relative positioning in global AI competitiveness as gaps remain in translating AI research into scalable applications, regulatory frameworks, and industry integration. Addressing these challenges requires coordinated policy interventions, enhanced industry-academia linkages, and strategic investment in human capital. Proactive measures—ranging from curriculum reform and skill development to fostering innovation ecosystems—can mitigate the risk of lagging in this pivotal era. This study examines the risk of India falling behind in the global AI revolution and highlights the importance of aligning the country's AI capabilities with international benchmarks, offering a strategic framework to maintain technological leadership while addressing the ethical, social, and economic dimensions of AI development.

Keywords: Artificial Intelligence, India, Technology, Policy, Research, and Ethics

1 METHODOLOGY

This paper adopts a qualitative and analytical methodology based on secondary research, drawing from government policy documents, industry reports, academic literature, and credible media sources to critically examine India's position in the global AI ecosystem. By employing comparative analysis with leading AI nations such as the United States and China, the study identifies structural gaps in talent, infrastructure, research, and regulation. This approach allows for a comprehensive evaluation of India's strengths and shortcomings without relying on primary data, situating the work within the tradition of critical policy and technology studies.

2 DELIBERATING STATEMENT OF THEME

The meeting between Sam Altman, CEO of OpenAI, and Ashwini Vaishnaw, India's IT Minister, on 5 February 2025, was more than a courtesy call—it symbolized both convergence of interests and a recognition of India's urgent need to strengthen its AI ecosystem. India is one of the fastest-growing AI markets, second only to the United States in OpenAI's user base, and the government has expressed its ambition to build the entire AI stack, from compute infrastructure like GPUs to foundational models and applications ([Business Today, 2025](#)). This highlights India's intent to move beyond being primarily an adopter of AI tools into becoming a producer of cutting-edge innovation. The meeting explicitly shows the promise of collaboration and the scale of India's digital economy, but it also implicitly acknowledged the gap between India's ambitions and the achievements of global leaders such as OpenAI ([NDTV, 2025](#)).

Despite a thriving startup ecosystem, government initiatives, and rising adoption across sectors, India still lags in critical areas like large-scale compute infrastructure, high-quality datasets, deep research labs, and venture capital for deep-tech AI (Analytics India Magazine, 2024; [IndiaAI, 2024](#)). Many Indian startups focus on fine-tuning or building applications on top of existing global

models rather than developing original, large-scale foundational models. Limited access to affordable GPUs, regulatory uncertainties, and brain drain add to the challenge ([The Outpost, 2024](#)). Thus, the Altman-Vaishnaw meeting can be read not as an admission of failure, but as an acknowledgement of the distance yet to be covered. India has the talent, demand, and policy vision to emerge as a significant AI player, but matching the stature of OpenAI will require sustained investment, global partnerships, and a long-term commitment to research and infrastructure.

India stands at a crossroads as the global Artificial Intelligence revolution accelerates. If the country fails to keep pace, it risks economic stagnation, widening inequality, and losing out in both innovation and global influence. To understand these risks properly, one must trace how India arrived here: the history of its tech revolution reveals both strengths to build upon and vulnerabilities that threaten to hold it back.

The roots of India's tech trajectory go back well before the internet boom. Some of the foundational work for computing in India was laid during the post-independence period when research institutes, state involvement, and a push for self-reliance in electronics were prominent ([Sharma, 2015](#), p.1-20). Early computers in government and university labs, state telecommunications systems, and government policies around electronics and computer import regulation all played roles in shaping the baseline capacity and human capital for what was to come ([Sharma, 2015](#), p.20). However, pervasive regulation, protectionism, and lack of scale constrained innovation and global competitiveness until reforms took root.

The turning point came with the economic liberalization of 1991. Facing a balance of payments crisis, India under Prime Minister P. V. Narasimha Rao and Finance Minister Manmohan Singh initiated sweeping reforms: removal of many licensing requirements, deregulation of interest rates, liberalization of foreign investment, and reduction of tariff and trade barriers (Hindu, 2014). These reforms facilitated trade growth, which rose from a modest share of GDP pre-1991 to much higher levels in ensuing decades, helped foreign firms enter, and allowed export-oriented software and services firms to scale (Times of India, 2023). Yet even this growth had trade-offs: much of the job creation in the post-liberalization era proved informal rather than formal, about 90% of jobs created over 22 years after liberalisation were informal ([Scroll, 2019](#)), which meant that many workers remained vulnerable.

By the late 1990s and early 2000s, India had established itself as a major destination for outsourcing, software services, and IT-enabled services (ITES). Educational institutions such as the IITs, engineering colleges, and the rise of private training institutions provided a large skilled workforce; policies such as Software Technology Parks, duty exemptions for hardware imports for exporters, and the advent of stable telecom infrastructure enabled firms to package and export digital services globally ([Sharma, 2015](#), p.150). This period showed India's ability to optimize its advantages: cost arbitrage, English language skills, and human capital. But these advantages, while powerful, were more about services rather than product innovation or control over underlying hardware and core infrastructure.

The AI revolution presents a qualitatively different challenge. AI demands not only software skills but also hardware capacity, such as semiconductors and GPUs, data infrastructure, regulatory regimes, and cutting-edge R&D. If India fails to build strength in those domains, there is risk of being a consumer or integrator of AI made elsewhere, rather than a creator. Already, reports suggest that India remains heavily reliant on imports for electronic components, including semiconductors, printed circuit boards and other hardware; nearly 80% of such inputs are imported, making the supply chain vulnerable to external shocks ([NITI Aayog report, 2024](#)). Additionally, the semiconductor sector faces a severe personnel shortage: estimates suggest that by 2027, India will be short by 250,000 to 300,000 semiconductor professionals in specialized areas ([Business Standard, 2024](#)).

Another risk is the large-scale displacement of jobs. Some reports suggest that ~68% of white-collar workers fear automation in the next five years, with about 40% fearing redundancy of current skills (Economic Survey, 2025; Down To Earth, 2025). Studies also warn that large numbers of middle and lower-income workers are most vulnerable, since many are currently employed in routine, repetitive or support roles, may find their jobs threatened by generative AI or related automation (Economic Survey, 2025).

India's history of rapid growth in services and outsourcing provided lessons in scaling labour and human capital, but left gaps in innovation, hardware sovereignty, and formal job creation. The risk now is that, without investment in deep-tech R&D, semiconductors, indigenous AI model development, and robust regulation for privacy, bias, and safety, India could fall behind nations that are aggressively investing in these areas, e.g. China, the US, and the European Union (Das, Mint, 2025). Moreover, the digital divide poses a serious internal

threat. Urban centres, elite engineering institutions, and the startup ecosystem have access to broadband, computing power, and funding; many rural or less privileged regions lag. If AI-driven innovation and productivity gains are concentrated in certain geographies and socio-economic groups, inequality could deepen sharply. The lessons from the early tech boom — that growth does not automatically produce inclusive development — are relevant again. For example, even during the post-liberalization period, a vast majority of job growth was informal, which meant that many did not gain stable wages, social protections, or predictable career paths ([Scroll, 2019](#)).

In terms of remedial pathways, history suggests some strategies that can help India avoid lagging. First, policy clarity and timely reform, as in 1991, are essential. Regulations for data protection, AI safety, standards for transparency, and liability for AI-related harms, etc, needed to be defined clearly. Second, AI requires not only software and human capital, but hardware like semiconductors and data centres, widespread connectivity, electricity, cooling, etc. Therefore, investment is a must. Third, human capital and curricula aligned with AI/ML/data science, reskilling mid-career workers, ensuring that skills do not get locked in those at risk of obsolescence, are needed. Fourth, promoting innovation and product development by encouraging firms that build core AI models, tools, hardware, and platforms would enhance India's potential as an AI leader. Finally, ensuring that the benefits of AI reach remote, rural, and underprivileged regions by bridging digital divides and addressing local language and cultural contexts is the need of the hour. India's tech revolution shows both what is possible when state, policy, human capital, and market forces align, and what gaps remain (Emphasize, author).

3 ENERGIZING TOGETHER: GOVERNMENT, INDUSTRY, AND ACADEMIA

The Indian government has indeed recognized the transformative potential of artificial intelligence, and initiatives like the IndiaAI Mission, with a ₹10,300 crore budget over five years, with nearly 18,700 GPUs planned for a national compute facility, demonstrate a strong policy commitment to positioning India as a serious AI player globally ([Press Information Bureau, 2025](#)). However, while the 2018 *National Strategy for Artificial Intelligence* laid out ambitious sectoral priorities—healthcare, agriculture, education, and smart cities—the

persistent gaps in foundational R&D, limited availability of high-quality datasets, and inadequate ecosystem support raise concerns about whether these investments will translate into global leadership or remain largely aspirational. Without parallel reforms in research capacity, semiconductor development, and regulatory frameworks, India risks becoming an adopter of AI applications rather than a leader in creating original technologies, especially compared to the rapid advances in the U.S., China, and the EU ([Press Information Bureau, 2025](#)).

While India's premier academic institutions, such as the IITs and other top universities, have expanded their AI curricula and research programs, and platforms like the National Olympiad in Artificial Intelligence (NOAI) aim to cultivate early talent, the broader picture remains far less encouraging. The scale and quality of AI research output from Indian universities still lag behind global leaders, with few publications in top-tier AI conferences and limited breakthroughs in foundational model development. Moreover, much of the academic focus remains theoretical or service-oriented, without the strong industry-academia linkages necessary to translate research into impactful innovation. These shortcomings highlight a structural weakness: India is producing a large number of graduates with AI exposure, but not enough high-quality researchers and innovators who can compete with counterparts in the U.S. or China. Without bold reforms in research funding, lab infrastructure, and incentives for deep-tech innovation, the current initiatives risk producing breadth without depth, leaving India vulnerable to technological dependence on foreign advances ([Das, 2025](#)).

India's fragmented nature of the research ecosystem, where funding is spread thin across institutions without building globally competitive AI research hubs. Even the IITs, despite their prestige, struggle with retaining top faculty and researchers, many of whom migrate abroad for better resources and collaboration opportunities. As a result, India's academic contributions often remain incremental rather than transformative, reinforcing the critique that current educational and talent-building initiatives, while commendable in intent, fall short of creating the innovation depth required for global AI leadership ([Rath, 2015](#), p.88).

Despite India's growing enthusiasm for AI adoption through startups and initiatives like the IndiaAI Datasets Platform (AIKosh), structural challenges continue to prevent the emergence of a global-standard AI company. Most Indian startups tend to focus on applied AI for local contexts—healthcare

diagnostics, agriculture monitoring, or language solutions—rather than developing foundational models or cutting-edge research at scale. This is partly due to limited access to advanced computing infrastructure, venture capital constraints for deep-tech, and the absence of a strong pipeline for high-quality R&D talent. While India has succeeded in producing service-based IT giants such as Infosys, TCS, and Wipro, no equivalent exists in AI, where global dominance requires massive capital expenditure, sustained research ecosystems, and a culture of risk-taking innovation that India's ecosystem has yet to foster ([Rath, 2015](#), p.91).

Collaborations between academic institutions and industry players, such as hackathons run by IISc and IBM, are positive but remain insufficient to bridge the innovation gap. Unlike the U.S. or China, where academic research often translates into world-class startups and unicorns in frontier technologies, India struggles with fragmented research output, low patenting rates, and weak commercialization pathways. Furthermore, issues like data privacy regulation, inadequate semiconductor capacity, and dependence on foreign hardware reinforce the perception that India risks being a follower rather than a leader in AI innovation. Unless systemic reforms address these gaps, India's AI ecosystem may remain strong in adoption and application but weak in producing globally competitive companies comparable to OpenAI, Anthropic, or DeepMind (NASSCOM, 2024).

India's IT advancement and AI quest highlight a striking paradox. While India has achieved global recognition as a hub for IT services and outsourced cybersecurity expertise, it continues to struggle with securing its own domestic digital infrastructure and advancing its own AI ecosystem. India's success abroad, where its professionals contribute significantly to the global AI journey and cyber industry, contrasts sharply with persistent weaknesses at home—ranging from inadequate national frameworks and underinvestment in indigenous capabilities to fragmented policies and poor coordination among institutions. This imbalance highlights how India's dependence on external technologies and lack of robust, homegrown AI models and mechanisms leave it vulnerable, despite its vast talent pool and global reputation.

The penetration of artificial intelligence in the Indian market demonstrates both measurable progress and substantial untapped potential. Recent industry analyses estimate that the Indian AI market will expand to approximately USD 17 billion by 2027, reflecting a compound annual growth rate (CAGR) of 25–35 per cent between 2024 and 2027 ([Reuters, 2024](#)). Within this trajectory, the

generative AI segment is valued at USD 1.02 billion in 2024 and is projected to grow to USD 8.3 billion by 2030, highlighting the rising adoption of AI-enabled solutions across sectors such as finance, healthcare, and e-commerce (Grand View Research, 2024). These figures emphasize that India is beginning to capture economic value from AI deployment, largely driven by enterprise adoption and increasing startup participation.

Nevertheless, the Indian AI ecosystem remains constrained by systemic challenges that inhibit full capitalization of its potential. Unlike global leaders such as the United States or China, India has yet to establish strong indigenous capacities in foundational model development, high-performance compute infrastructure, and semiconductor manufacturing. A large proportion of AI innovation remains concentrated in applications and service-driven models rather than deep research or hardware sovereignty. Further, regulatory uncertainty and limited access to curated, high-quality datasets restrict the scalability of indigenous AI solutions. Consequently, while projections suggest robust growth, the structural limitations indicate that India risks under-realizing the transformative economic value of AI unless it deepens investments in research, infrastructure, and regulatory frameworks to convert its demographic and digital advantages into global competitiveness.

Table 1. Major AI platforms — India Presence & Paid-User Information.

Platform	Estimated India Users / Share	Paid Users in India	Global Paid / Customers (Public)	Source (India presence)	Source (Global paid)
OpenAI — ChatGPT	India is OpenAI's 2nd-largest market; large user base but exact India figures undisclosed.	Not disclosed.	Approx. 15.5M+ paid subscribers globally (2024 est.).	Business Today, 2025 ; OpenAI reports.	The Information, 2024

Platform	Estimated India Users / Share	Paid Users in India	Global Paid / Customers (Public)	Source (India presence)	Source (Global paid)
Google — Gemini / Bard	India $\approx$ 9–10% of global traffic; often 2nd-largest market.	Not disclosed (mostly free access).	$\approx$ 140M monthly visits globally (2024–25).	Meetanshi / Scoop (2025)	Meetanshi / traffic analyses
Microsoft — Copilot / GitHub Copilot	Strong enterprise adoption in India; exact figures not published.	Not disclosed.	GitHub Copilot: 20M+ users globally (2025); Copilot integrated into MS products has tens of millions users.	Microsoft reports (2024–25)	TechCrunch (2025); Microsoft AR24
Anthropic — Claude	Enterprise users in India but no published counts.	Not disclosed.	No public paid-subscriber numbers; case studies available.	Anthropic customer pages	Anthropic reports

Platform	Estimated India Users / Share	Paid Users in India	Global Paid / Customers (Public)	Source (India presence)	Source (Global paid)
Meta — Meta AI / LLaMA	Meta's consumer apps have huge India bases; LLaMA usage not published.	N/A (integrated in free apps).	Billions of users across Meta apps; no AI-specific figure.	Meta reports / product stats	Meta reports

Important caveats: country-level paid-subscriber numbers are rarely disclosed by these companies, so most India-specific ‘paid user’ entries are not publicly available and marked accordingly. The table uses traffic-share and corporate reporting sources.

In a sensational report, OpenAI announced that it had surpassed 1 million paid business users of ChatGPT in September 2024, reflecting growing enterprise adoption of its AI tools. In parallel, the free tier continues to attract hundreds of millions of monthly users worldwide, highlighting both widespread popularity and the company's successful penetration across consumer and business segments (Reuters, 2025). Since the launch of ChatGPT in late 2022, India has seen very strong adoption of AI tools — in usage, attention, and downloads — but almost no reliable public data on how much Indian consumers are paying in aggregate for AI subscriptions, whether ChatGPT Premium/Plus/Enterprise, or competing offerings. For example, SensorTower data show that Indian users have spent about US\$8 million on ChatGPT subscriptions via in-app payments since 2023, excluding web-based payments. In contrast, users in the U.S. have spent ~US\$330 million over the same period. The gap between adoption and monetization suggests not only that pricing and ability/willingness to pay are major barriers, but also that much of India's paid-use, if any large-scale exists, is not captured in regular commercial reporting or analytics, particularly because many payments may be done via non-app or enterprise channels.

A report by International Data Corporation and Qlik finds that spending by Indian enterprises on artificial intelligence is projected to grow at about 35% annually, reaching US\$9.2 billion by 2028 ([Economic Times, 2025](#)). Key findings in this report include that 36% of Indian organizations are currently using generative AI, and another 46% plan to invest in it over the next 12-24 months.

This opacity has broader implications. It constrains the ability of researchers, policy makers, and industry to understand the real economic contribution of AI subscriptions, and to shape regulation, taxation, and support policies accordingly. The Reserve Bank of India (RBI) and banks presumably have data on foreign remittances or recurring payments to overseas digital service providers (DSPs), which would include many AI subscription services (India Express, 2019). The RBI has clarified that there is no restriction on payment processing abroad, and it has reiterated that all payment data must be stored within India to ensure regulatory access and security. This stance comes amid ongoing government discussions on tightening data localization rules. Government officials argue that local data storage would strengthen investigative capacity and oversight, even as industry stakeholders worry about compliance costs and operational hurdles (Indian Express, 2024). However, such data is either not disaggregated in public reports or access is restricted. This lack of transparency may lead to an underestimation of how much value is flowing from Indian consumers to global tech giants, which in turn can limit conversations about fair revenue sharing, digital sovereignty, and consumer protection (Indian Express, 2024).

From a critical standpoint, this situation can be conceptualized as a form of economic asymmetry. The pattern of consumption of AI tools, models, and services is ahead of the pattern of domestically generated information and policy response. India is importing or subscribing to AI capabilities where local actors do not yet provide equivalent services; payments are made overseas, but the financial flows, taxation, regulation, and reward for creators are uneven. Unless India's financial regulators, tax authorities, and digital oversight bodies mandate clearer reporting, or unless private sector firms publish India-specific subscription/revenue data, this gap will persist. This may contribute to a kind of 'digital economic imperialism', where value accrues mostly to firms based elsewhere, while Indian users bear the costs but receive limited lifelong value or bargaining power.

To get a sense of AI users in India, I gathered available survey/traffic analytics data + indirect sources, then built a ‘best estimate’ table showing AI user growth in India since 2022 after the launch of ChatGPT. The caveats about this process are that many numbers are approximate; ‘users’ includes free + paid. The platform breakdowns are based on surveys and traffic share; paid vs free is largely unknown.

Table 2. AI Users in India, by Platform (2022–2025).

Year	Total AI Users in India*	ChatGPT Users**	Other Platforms / Key Competitors (approx)	Notes / Data Sources
2022	~10-15 million	~5-8 million	Few users on newer platforms (Perplexity, Bard, etc.)	Based on early adoption curves, download stats, plus app analytics reports, 2022 was an early stage after ChatGPT's launch.
2023	~35-45 million	~25-30 million	~10-15 million across “others” (Google Bard / Gemini, Perplexity, Co-Pilot, etc.)	Growth is indicated by increasing app downloads, web traffic, and survey-based awareness.

Year	Total AI Users in India*	ChatGPT Users**	Other Platforms / Key Competitors (approx)	Notes / Data Sources
2024	~100-130 million	~70-80 million	~30-50 million using alternatives (Gemini, Bard, Perplexity, Co-Pilot, etc.)	Supported by reports such as IndiaAI app download share (~21% of downloads), LocalCircles surveys showing ~50% of surveyed internet users using AI platforms, and ChatGPT India MAU share (estimates ~13-14%) in Mary Meeker / similar reports.
2025 (early/mid)	~150-180 million	~100-120 million	~50-70 million “others”	Based on continued growth, rising internet penetration, and increased use in enterprise and consumer segments (estimates from surveys + traffic trends).

Note: [↗](#) * ‘Total AI Users’ here means monthly active users (MAUs) of any major conversational/generative AI platform (free + paid) in India.

Note: [↗](#) ** ChatGPT Users refers to those using ChatGPT products (chat.openai.com, apps) on free or subscription tiers.

It has been observed that ChatGPT appears to have the lion's share of users in India among generative AI chat platforms, especially active from late 2022 to 2023. Platforms like Google's Bard / Gemini, Perplexity, Microsoft's Co-Pilot, etc., are growing, especially among tech-savvy users and enterprise users. Very little reliable public data exists for how many of the above users are paying vs using free tiers. As noted earlier, companies do not typically disclose country-level paid subscriber numbers. These estimates are built from surveys, app

downloads & traffic analytics, and news reports — so margins of error are large. Also, definitions of ‘active user’, ‘AI platform’, and ‘free vs paid’ vary across sources.

4 INDIA: RISKS OF LAGGING IN AI REVOLUTION

India's aspirations to become a global leader in AI face significant risks if structural challenges are not addressed. Despite rapid progress in policy frameworks and infrastructure development, the nation struggles with gaps in skilled manpower and technological readiness. While the AI sector could generate 2.3 million jobs by 2027, India is projected to produce only about 1.2 million qualified professionals (Goldman [Sachs, 2025](#)). This shortfall suggests that without targeted investment in advanced AI education and training, India risks missing out on a substantial share of high-value opportunities in the global AI economy.

In addition to the talent gap, infrastructural disparities remain a pressing issue. Urban regions benefit from growing investments in data centres and AI compute facilities, while rural areas continue to face unreliable internet access and limited computational resources. This uneven distribution restricts the democratization of AI, concentrating benefits in metropolitan hubs while excluding much of the population from participating in and contributing to the AI revolution. Bridging this divide is essential for inclusive growth and ensuring AI applications address India's diverse socio-economic challenges.

Regulatory preparedness also poses a major hurdle. The pace of AI advancement is outstripping the evolution of legal and ethical frameworks, leaving gaps in governance around data privacy, bias, accountability, and misuse. Without comprehensive regulation, India risks creating an environment where innovation proceeds unchecked at the expense of social trust and ethical safeguards. This regulatory lag could undermine both domestic adoption and international credibility in the AI sector (Goldman [Sachs, 2025](#)).

Global competition amplifies these vulnerabilities, especially when countries like the United States and China are heavily investing in deep-tech R&D, AI infrastructure, and cross-border collaborations; they are consolidating their leadership positions. India, by contrast, risks being confined to the role of a service provider rather than an innovation leader if it does not scale up investment and foster indigenous capabilities. Falling behind in the global AI

race could result not only in economic opportunity costs but also in strategic dependence on foreign technologies, thereby reducing India's autonomy in shaping its digital future.

There is an urgent need for India's education system to adapt to the transformative impact of artificial intelligence. Despite AI's rapid adoption across industries worldwide, Indian institutions have been slow in integrating it into classrooms, curricula, and research practices. For instance, even in recent recruitment notices, few universities or colleges explicitly mention AI skills, which reveals a systemic lack of recognition of AI's growing relevance. This complacency risks leaving Indian students and faculty ill-prepared for the demands of an AI-driven global economy ([Singh, 2025a](#)).

This is pertinent especially in the field of mathematics and applied sciences, domains that have already seen groundbreaking transformations due to AI. Neural-network-based techniques are being used internationally to solve differential equations and model complex real-world phenomena, yet Indian educators remain largely unaware of or resistant to these tools. By not engaging with such advancements, India risks maintaining outdated pedagogical practices and failing to nurture students capable of applying cutting-edge methods. Embracing AI is not merely about adopting new technologies, but also about reshaping the intellectual approach of teachers and researchers to align with contemporary global standards.

A comprehensive cultural shift in the education sector is needed, and faculty members, schoolteachers, and higher-education leaders must take responsibility for equipping themselves with AI knowledge and fostering curiosity among students. India cannot afford to let AI be treated as a passing trend; instead, it must be embedded into the foundations of teaching, learning, and research. Without this proactive approach, the country risks allowing AI to bypass its education system entirely, leaving its young population dependent on foreign innovations rather than contributing to them meaningfully ([Singh, 2025a](#)).

The present scenario is that India is missing the crucial window to establish itself as a leader in artificial intelligence. While countries like the United States and China are investing heavily in foundational AI research and creating advanced models capable of solving doctoral-level problems, India's academic and policy environments remain sluggish. India, as a nation, is primarily consuming AI technologies developed abroad rather than building indigenous capabilities. This trend reflects a deeper problem. India is approaching AI

passively, as a user and market, instead of positioning itself as a creator of global-standard solutions ([Singh, 2025b](#)).

The rigidity of India's education system continues to rely on outdated curricula and methods. While AI is transforming fields like mathematical modelling, data analysis, and reasoning worldwide, Indian institutions have not meaningfully incorporated these shifts into their programs. The absence of dynamic engagement with AI in universities and research centres indicates a lack of vision, leaving students underprepared for a future dominated by advanced computational tools. Such inertia risks widening the talent gap and diminishing India's competitiveness in both research and industry.

India needed to be cautioned that the cost of inaction could be steep. By failing to catch the AI bus at the right moment, India risks long-term economic, intellectual, and strategic disadvantages. Instead of being at the forefront of innovation, India could become permanently dependent on AI systems developed overseas, undermining its sovereignty in technology and knowledge production. The warning is clear: unless India invests seriously in AI research, education reform, and innovation ecosystems, it may pay a heavy price by forfeiting its chance to lead in one of the most critical technological revolutions of the century ([Singh, 2025b](#)).

5 CONCLUSION

India's engagement with artificial intelligence reveals an uncomfortable paradox: while it boasts one of the largest user bases of AI tools globally, it has yet to establish itself as a producer of foundational innovation. The country remains trapped in a consumption-driven model, where global platforms extract value from Indian users without corresponding domestic capacity in research, infrastructure, or commercialization. This reliance creates structural dependency on foreign technologies, a form of digital subordination that echoes earlier failures in hardware sovereignty and industrial innovation. Such dependence threatens not only economic opportunities but also long-term technological autonomy.

The repeated shortcomings of India's academic and industrial ecosystem underscore systemic inertia. Universities continue to emphasize rote learning and service-oriented applications rather than fostering high-quality research in deep-tech, while the startup landscape prioritizes quick monetization over long-

term innovation. This shallow integration of AI into education and research has already created a widening talent gap, as highlighted by persistent shortages of advanced AI professionals and semiconductor specialists. Without addressing these structural weaknesses, India risks being confined to a secondary role in the AI revolution—an integrator and consumer, rather than a creator of transformative technologies.

Ultimately, India's hesitation to reform at the depth and scale required exposes it to severe consequences. Countries like the United States and China are accelerating investments in AI infrastructure, patents, and frontier research, consolidating their dominance. India, by contrast, appears content with incremental progress and symbolic announcements. Unless the nation embraces bold reforms in research funding, regulatory clarity, and education, it risks forfeiting its demographic and digital advantages. The cost of this inaction will be steep: the erosion of global competitiveness, heightened inequality, and permanent dependence on foreign innovation to shape its digital future.

REFERENCES

- Analytics India Magazine, *Small Successes Mask Fundamental Gaps in India's AI Ecosystem*, 2024, retrieved from <https://analyticsindiamag.com/ai-startups/small-successes-mask-fundamental-gaps-in-indias-ai-ecosystem/>
- Business Standard, TeamLease Degree Apprenticeship report, *India's Chip Industry to Face Shortage of 300,000 professionals by 2027*, 11 June 2024.↵
- Business Today, *India Incredibly Important Market for AI: OpenAI CEO Sam Altman Meets Ashwini Vaishnaw on India Visit*, 1 February 2025, retrieved from <https://www.businesstoday.in/technology/news/story/india-incredibly-important-market-for-ai-openai-ceo-sam-altman-meets-ashwini-vaishnaw-on-india-visit-463538-2025-02-05>↵
- Chandrasekaran N., and R. Purushothaman, *Bridgital Nation: Solving Technology's People Problem*, New Delhi: Penguin Allen Lane, 2019. L. V. Subramaniam, *Quantum Nation: India's Leap Into the Future*, New Delhi: Cleverfox Publishing, 2022.
- China Briefing, *India's AI Infrastructure and Emerging Market Leadership*, 2025, retrieved from <https://www.china-briefing.com/china-outbound-news/indias-ai-infrastructure-and-emerging-market-leadership-an-outlook>.

Dinesh Singh, 'Wake up to the importance of AI in education', New Indian Express, 1 February 2025a.

Dinesh Singh, India is letting the AI revolution bypass the country. It may have to pay a heavy price, The Print, 9 July 2025b. [↵](#)

Dinesh Singh, *India is Letting the AI Revolution Bypass the Country. It May Have to Pay a Heavy Price*, 9 July 2025a, The Print, retrieved from <https://theprint.in/opinion/india-ai-revolution-bypass-country-pay-heavy-price/2683999/>. [↵](#)

Dinesh Singh, *Wake up to the importance of AI in education*. The New Indian Express, 1 February 2025b, retrieved from <https://www.newindianexpress.com/magazine/voices/2025/Feb/01/wake-up-to-the-importance-of-ai-in-education>.

Economic Survey of India, *Chapter on AI Evolution and Job Automation Risks*, 2025.

Economic Times, AI spending in India set to reach \$9.2 billion by 2028: Report, 16 April 2025. [↵](#)

Goldman Sachs, The future of AI talent and opportunities in India, Goldman Sachs Research, 2025. [↵](#)

India Today, *OpenAI's Sam Altman meets with IT Minister Ashwini Vaishnaw to discuss India AI roadmap*, 5 February 2025, retrieved from <https://www.indiatoday.in/business/story/openai-sam-altman-meets-with-it-minister-ashwini-vaishnaw-to-discuss-india-ai-roadmap-2675222-2025-02-05>

IndiaAI, *Generative AI startup landscape in India*, 2024, retrieved from <https://indiaai.gov.in/research-reports/generative-ai-startup-landscape-in-india> [↵](#)

Indian Express, No bar on payment processing abroad but data should be stored in India only, says RBI, 27 June 2019.

Mehta H.S., *The Maverick Effect: The Inside Story of India's IT Revolution*, New Delhi: HarperCollins, 2022.

Microsoft, *Microsoft Announces \$3 Billion Investment Over Two Years in India Cloud and AI Infrastructure*, 2025, retrieved from <https://news.microsoft.com/source/asia/2025/01/07/microsoft-announces-us-3bn-investment-over-two-years-in-india-cloud-and-ai-infrastructure-to-accelerate-adoption-of-ai-skilling-and-innovation>.

NDTV, *Union Minister Ashwini Vaishnaw has "super cool" discussion with OpenAI CEO Sam Altman*, 5 February 2025, retrieved from

<https://www.ndtv.com/india-news/union-minister-ashwini-vaishnaw-has-super-cool-discussion-with-openai-ceo-sam-altman-7639333>

NITI Aayog report, India Heavily Reliant on Imports for Component Manufacturing, 2024.

Press Information Bureau, *India's AI Revolution*, 2025, retrieved from <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2108810>

Reuters, OpenAI hits more than 1 million paid business users, Indian Express, September 7, 2024.

Saroj Kumar Rath (Ed.), *India's cyber security*, 'India's Cyber Insecurity: Won the World, Defeated at Home', New Delhi: Synergy Books India, 2015.

Sauvik Das, India is missing the core elements needed to realise the AI dream, *Mint*, 11 April 2025, retrieved from <https://www.livemint.com/news/india-ai-lack-of-innovation-dream-er-d-research-spending-patents-generative-ai-china-samsung-sarvam-ganai-11750336848932.html>.

Scroll, Data check: 90% of jobs created in two decades post liberalisation were in the informal sector, 2019 *Scroll.in*.

Sharma D.C., *The Outsourcer: The Story of India's IT Revolution*, Boston: MIT Press, 2015.

Sunil Khilnani, *The Idea of India*, New York: Farrar, Straus and Giroux, 1997.

The Outpost, *Indian AI startup ecosystem: Funding dip, deal volume rise, and VC optimism*, 2024, retrieved from <https://theoutpost.ai/news-story/indian-ai-startup-ecosystem-funding-dip-deal-volume-rise-and-vc-optimism-9893>

Securitizing Artificial Intelligence Through a Human Security Lens: Historical Trajectories, Emerging Risks, and Normative Futures

Alp Cenk Arslan and Yasemin Kaymaz

Turkish National Police Academy, Ankara, Turkey

DOI: [10.1201/9781042018116-10](https://doi.org/10.1201/9781042018116-10)

ABSTRACT: This chapter examines the securitization of artificial intelligence (AI) through a human security lens, critiquing dominant state-centric paradigms that prioritize geopolitical dominance and existential risks over individual vulnerabilities. Drawing on the Copenhagen School's securitization theory (Buzan *et al.*, 1998) and the UNDP's seven pillars of human security (United Nations Development Programme, 1994), it reveals AI as a dual-use technology. A tool for state fortification that paradoxically erodes personal, economic, and political securities via surveillance, bias, and exploitation. Tracing historical roots from Cold War cybernetics and Project Cybersyn's democratic promise to post-9/11 Total Information Awareness and the AI “Revolution in Military Affairs,” the analysis exposes how narratives of arms races and superintelligence (Ord, 2020) mask immediate harms, including algorithmic paranoia in gig work (Alacovska *et al.*, 2025), biased healthcare (Federspiel *et al.*, 2023), environmental footprints, and digital colonialism extracting Global South data (Nothias, 2025). Evaluating regulatory responses, the EU AI Act's risk tiers (Powell, 2024), African Union's developmental strategy (African Union Commission, 2024), and UNESCO's ethical framework (UNESCO, 2021), it advocates desecuritization of routine biases and securitization of

data rights. In conclusion, a human-centric architecture demands transparency, sovereignty, and audits to transform AI from a fragility vector to a shared dignity scaffold, fostering “freedom from fear and want.”

Keywords: Human Security, Securitization Theory, AI Governance, and Global Ethics

1 METHODOLOGY

This chapter employs a qualitative, theory-driven methodology combining historical analysis with normative policy critique. It applies securitization theory and the human security framework as analytical lenses to trace shifting AI security narratives across historical and contemporary contexts. Comparative examination of global regulatory regimes and illustrative case studies is used to assess emerging risks, while normative reasoning is deployed to propose human-centric futures for AI governance.

2 INTRODUCTION: THE ALGORITHMIC SECURITY PARADOX

The advent of the twenty-first century has been irrevocably marked by the ascendancy of Artificial Intelligence (AI) as a foundational technology of global order. No longer confined to the speculative realms of science fiction or the isolated laboratories of computer science departments, AI has permeated the granular texture of everyday life, mediating our access to information, credit, healthcare, and justice. Simultaneously, it has ascended to the highest levels of statecraft, becoming the central object of a new great power competition that frames algorithmic superiority as the *sine qua non* of national survival. The concept of “security” is being radically renegotiated in this volatile landscape. The discourse surrounding AI is increasingly dominated by a lexicon of existential fears, “arms races”, “kill switches,” “singularities,” and “strategic offsets”. Yet, this prevailing state-centric securitization, focused on the preservation of sovereignty and the

balance of power, often operates as a mechanism of concealment. By directing the collective gaze toward hypothetical future catastrophes or abstract geopolitical rivalries, it renders invisible the immediate, material insecurities inflicted upon individuals and communities by the deployment of these very technologies.

This chapter asserts that the current securitization of AI is fundamentally misaligned. It argues for a normative and analytical pivot from the “State Security” paradigm, prioritizing hard power assets of the nation-state such as territorial integrity and military capacity, to a “Human Security” paradigm, prioritizing the safety, dignity, and agency of the individual. Using the theoretical architecture provided by the Copenhagen School of Security Studies, specifically the work of Barry Buzan and Ole Wæver, and the United Nations Development Programme's (UNDP) human security framework, this analysis seeks to map the complex topography of AI risks.

The central thesis emerging from this research is that AI functions as a “dual-use” technology not merely in the military sense (civilian vs. weapon), but in an ontological sense. It is simultaneously a tool of security provision and a vector of profound insecurity ([Lupovici, 2021](#)). For the state, AI surveillance may enhance “security” by predicting unrest for the citizen, that same surveillance erodes “personal security” by dismantling privacy and enabling authoritarian control. For the corporation, algorithmic management increases “economic security” through efficiency for the gig worker, it generates “algorithmic insecurity” and psychological paranoia ([Alacovska et al., 2025](#)). This divergence creates a security paradox where the fortification of the system often necessitates the vulnerability of the subject.

To understand how AI is being governed, it must first be understood how it is being spoken about. Securitization theory teaches us that security is not an objective state of affairs but a “speech act”, a performative utterance that moves an issue from the realm of normal politics into the realm of emergency, justifying extraordinary measures ([Buzan et al., 1998](#)). When a “securitizing actor” (such as a defense minister or a tech CEO) frames AI as an “existential threat” (the referent object being humanity or the state), they are attempting to bypass standard democratic deliberation ([Zeng, 2021](#)).

This chapter dissects these speech acts, revealing a contest between competing narratives. On one side, the “Existential Risk” (X-Risk)

community frames AI as a future super-intelligence that could extinguish humanity, a narrative that ironically empowers the very tech companies building these systems to demand regulatory capture and immense resources ([Ord, 2020](#)). On the other side, civil society actors and scholars from the Global South attempt to securitize the present, framing the immediate harms of bias, extraction, and labor exploitation as the true emergency ([Milmo & Stacey, 2023](#)).

This analysis is exhaustive in its scope, traversing the historical, theoretical, and empirical dimensions of the AI-security nexus. We trace the genealogy of AI securitization from the cybernetic dreams of the Cold War and the socialist experiments of Project Cybersyn to the post-9/11 surveillance architectures of Total Information Awareness ([Glickman, 2024](#)). We conduct a forensic examination of how AI impacts the seven pillars of human security defined by the 1994 UNDP Human Development Report, economic, food, health, environmental, personal, community, and political security ([United Nations Development Programme, 1994](#)). We evaluate the emerging regulatory landscapes, contrasting the risk-based approach of the European Union's AI Act with the developmental focus of the African Union's AI Strategy and the ethical frameworks of UNESCO ([Powell, 2024](#)). By weaving together these disparate threads, this chapter aims to construct a “Human Security Lens” for AI, a diagnostic tool that allows policymakers to see beyond the hype of the “extreme existential risk scenarios” and address the quiet, systemic violence of the algorithm.

While existing literature extensively covers AI through either state-centric geopolitical lenses or abstract existential risks, this chapter offers a novel contribution by conceptualizing the vertical security dilemma. It is the moving beyond horizontal state-to-state rivalries to analyze the tension between state fortification and individual vulnerability. The second contribution is the integrating digital colonialism into securitization theory. Examining how data extraction from the Global South functions as a form of structural insecurity is a significant contribution in this chapter. Finally, proposing desecuritization as governance is one of the most prominent contributions. Arguing for the removal of routine algorithmic harms from “emergency” rhetoric to enable democratic deliberation and data rights should be conducted in the analysis of the chapter.

3 THEORETICAL FRAMEWORK: THE GRAMMAR OF ALGORITHMIC SECURITY

The analysis of AI requires a robust theoretical apparatus capable of bridging the gap between high politics and everyday life. The integration of Securitization Theory with the Human Security Paradigm provides this necessary bridge, allowing us to analyze AI not just as a technology, but as a political process of threat construction.

Originating in the work of Barry Buzan, Ole Wæver, and Jaap de Wilde, the Copenhagen School revolutionized security studies by shifting the focus from “what is a threat?” to “how is a threat constructed?” ([Buzan *et al.*, 1998](#)). This constructivist approach is vital for understanding AI, a technology whose risks are often intangible, probabilistic, or futuristic.

The mechanism of securitization involves distinct components. These are the securitizing actor, the referent object, the securitizing move, and finally the audience. In the AI domain, the securitizing actor role is occupied by a diverse cast. State actors (for instance, the US Department of Defense, the Chinese Communist Party) frame AI as a matter of geopolitical dominance. Ideally, they are joined by “functional actors”, tech CEOs and researchers, who possess the technical authority to define the boundaries of the possible ([Zeng, 2021](#)). The referent object is the entity that is purportedly threatened. In traditional discourse, this is the State. However, in the “AI Safety” discourse, the referent object is often “Humanity” itself, or “Civilization.” In the human security discourse, the referent object is the individual or the community ([Zeng, 2021](#)). The securitizing move is the speech act itself, the declaration that “unless we act now, X will destroy Y.” For AI, this move often takes the form of “arms race” rhetoric (“We must develop autonomous weapons before our adversaries do”) or “alignment” rhetoric (“We must solve control problems before AI becomes superintelligent”) ([Zeng, 2021](#)). A securitizing move is only successful if the audience accepts it. In the case of AI, the audience is dual. The general public, which is fed narratives of sci-fi dystopia, and the policy elite, which is fed narratives of economic competitiveness ([Zeng, 2022](#)).

The Copenhagen School also introduces the concept of “Macrosecritization,” which refers to the framing of threats that

encompass the entire globe or humanity as a whole ([Karyotis & Paterson, 2025](#)). The climate crisis is the paradigmatic example, but AI is rapidly being positioned as a second macrosecuritization event. The danger, as noted by Buzan and Wæver, is that successful securitization moves an issue out of “normal politics”, where debate and compromise are possible, into “emergency politics,” characterized by secrecy, executive power, and the suspension of rights ([Buzan et al., 1998](#)).

While the Copenhagen School focuses on speech, the Paris School (associated with Didier Bigo) focuses on practice ([Naghibeyli, 2025](#)). This lens is particularly relevant for AI. Much of the securitization of AI does not happen through dramatic speeches by presidents, but through the quiet, technocratic implementation of algorithms by bureaucrats and police officers.

When a border agency deploys a biometric identification system, or a welfare agency deploys a fraud-detection algorithm, they are engaging in “securitization by design.” The algorithm itself becomes the securitizing actor, constantly scanning for “risk” and “abnormality” ([Rinaldi & Teo, 2025](#)). This is a process of “risk-based securitization,” where the logic is not defense against an enemy, but the management of probable futures ([Ulloa Ruiz & Bas Graells, 2025](#)). The threat is not a hostile army, but a statistical deviation.

Against the state-centric obsessions of Realism, the Human Security paradigm offers a normative corrective. First articulated in the UNDP's 1994 Human Development Report, it reorients security around the individual ([United Nations Development Programme, 1994](#)). It argues that for the average person, “security” is not defined by the absence of foreign invasion, but by the presence of food, health, employment, and protection from crime and repression.

The seven pillars of human security serve as the analytical backbone of this chapter. The seven pillars are economic security, food security, health security, environmental security, personal security, community security, and political security. If we define each pillars in terms of AI and autonomous weapons, we should indicate as follows ([Koundouri et al., 2025](#));

Economic Security: The assurance of a basic income, now threatened by automation and the precarity of the gig economy.

Food Security: Physical and economic access to food, now mediated by global supply chains and precision agriculture algorithms.

Health Security: Protection from disease, now dependent on algorithmic diagnostics and resource allocation models.

Environmental Security: Protection from ecological degradation, now entangled with the massive carbon and water footprint of AI compute.

Personal Security: Protection from physical violence, now threatened by autonomous weapons and predictive policing.

Community Security: The preservation of group identity, now eroded by digital polarization and the loss of traditional social structures.

Political Security: The enjoyment of civil rights, now under siege from digital surveillance and disinformation.

Bridging these theories reveals the “Dual-Use Security Dilemma” inherent in AI ([Lupovici, 2021](#)). In international relations, the security dilemma describes a situation where one state's attempt to increase its security (by building AI weapons) decreases the security of others, leading to a spiral of escalation.

However, in the context of AI, this dilemma operates vertically between the state and the citizen. The tools used to secure the state, mass surveillance, data mining, border control algorithms, are the very tools that generate insecurity for the individual. The “success” of a state securitization move (achieving “Total Information Awareness”) often represents a “failure” of human security (the loss of privacy and liberty) ([Karyotis & Paterson, 2025](#)).

Furthermore, the concept of “Desecuritization” becomes crucial. If securitization causes the suspension of rights, then the goal of a human-centric approach should often be desecuritization, returning issues like migration or social welfare back to the realm of normal political debate, rather than managing them through emergency security algorithms ([United Nations Development Programme, 1994](#)). Conversely, some issues, such as the data rights of the Global South or the labor conditions of data workers, may need to be securitized (framed as urgent threats) to garner the necessary political attention ([Floyd, 2021](#)).

4 HISTORICAL TRAJECTORIES: THE GENEALOGY OF ALGORITHMIC SECURITY

The current “AI revolution” is not a sudden rupture but the latest iteration of a long historical process of integrating computation with control. To understand the present securitization of AI, we must excavate its roots in the Cold War, the cybernetic movement, and the post-9/11 surveillance state. The intellectual lineage of AI traces back to the mid-20th century discipline of cybernetics, pioneered by Norbert Wiener. Cybernetics envisioned the world, biological, mechanical, and social, as a system of information feedback loops. This worldview was deeply intertwined with the Cold War security state. The development of early computing was driven by the imperatives of nuclear ballistics, cryptography, and command-and-control systems ([Glickman, 2024](#)).

During this era, the US military adopted the “Offset Strategy” in the 1970s. Facing the Soviet Union's numerical superiority in conventional forces, the Pentagon sought a qualitative advantage through superior technology, specifically microelectronics and precision-guided munitions ([Glickman, 2024](#)). This was the first “offset.” It established the structural relationship between the defense establishment and the tech sector (Silicon Valley) that persists today. The logic was clear. National security depends on computational superiority.

However, this era also contained alternative futures. In Chile, under the socialist government of Salvador Allende (1970-1973), a different vision of algorithmic governance emerged. Project Cybersyn (Sistema Synco) ([Tucker, 2025](#)). Spearheaded by British cybernetician Stafford Beer, Cybersyn aimed to manage the national economy not through Soviet-style top-down command, but through a decentralized network of telex machines (Cybernet) connected to a mainframe. The system included a futuristic “Opsroom” where decision-makers could view real-time economic data ([Morozov & Higgins, 2023](#)).

Unlike the surveillance models that would follow, Cybersyn was designed with a “Viable System Model” that emphasized worker participation and autonomy ([Herrera, 2022](#)). It represented a “human security” approach to cybernetics, using technology to ensure economic

stability and equitable distribution ([Joseph, 2023](#)). The project was destroyed following the CIA-backed coup by Augusto Pinochet in 1973, marking a foreclosure of this democratic path and the triumph of the neoliberal, militarized model of computing ([Morozov & Higgins, 2023](#)).

The terrorist attacks of September 11, 2001, served as a catalyst for the aggressive securitization of data. The failure of intelligence agencies to “connect the dots” led to a paradigm shift from targeted surveillance to mass data collection.

This shift was embodied in the Total Information Awareness (TIA) program, launched by DARPA and led by Admiral John Poindexter. TIA's explicit goal was to create a “totalizing surveillance perspective” of the populace by aggregating data from financial records, travel manifests, medical records, and communications. The logic was “predictive policing” on a global scale, using algorithms to identify patterns indicative of terrorist planning before the event occurred ([Murray, 2010](#)).

Although TIA was publicly defunded by Congress in 2003 due to civil liberties concerns, its architecture did not disappear. It was merely compartmentalized and privatized. Programs with codenames like “Basketball” and “Topsail” continued the work within the NSA. The “logic” of TIA, that security requires the total ingestion of the world's data, became the foundational logic of the modern surveillance state and the business model of “Surveillance Capitalism.”

The genealogy of AI securitization reveals a persistent drive to integrate computational power with state control, oscillating between democratic promises and surveillance imperatives. Project Cybersyn (1970-1973) briefly envisioned a human-centric cybernetic governance aimed at economic stability and worker autonomy. However, this alternative was largely eclipsed by the post-9/11 paradigm of TIA, which prioritized mass data collection to achieve a totalizing surveillance perspective. This shift institutionalized the logic that national security requires the total ingestion of civilian data, setting the foundation for modern surveillance capitalism. This historical trajectory illustrates that the securitization of data is not a modern anomaly but a continuation of the state's drive for computational superiority, which sets the stage for today's human security tensions.

The cultural response to this securitization is exemplified by artist Hasan Elahi. After being erroneously investigated by the FBI post-9/11, Elahi

launched the project “Tracking Transience,” where he voluntarily surveilled himself, posting his location and financial data online continuously ([Elahi et al., 2022](#)). This act of “sousveillance” (watching from below) highlighted the absurdity and intrusiveness of the state's securitizing gaze.

In the 1990s, military theorists debated the “Revolution in Military Affairs” (RMA), arguing that information technologies would lift the “fog of war” and ensure rapid, decisive victories ([Wells, 2025](#)). The swift victory in the Gulf War (1991) seemed to validate this.

Today, this discourse has been resurrected as the “AI-RMA.” Proponents argue that the integration of AI into military systems, autonomous drones, decision-support systems, and cyber-warfare, represents a new discontinuity in warfare ([Daniels, 2022](#)). The current narrative frames AI not just as a force multiplier but as the decisive factor in future conflict. This drives the “arms race” mentality, where the development of Lethal Autonomous Weapons Systems (LAWS) is justified as a necessary response to the capabilities of adversaries like China and Russia ([Acheson, 2025](#)).

This historical trajectory reveals a consistent pattern. Technology is developed under the banner of “national security,” legitimized by the threat of an external enemy (Soviets, Terrorists, Great Power Rivals), and then integrated into the infrastructure of civilian control.

5 STATE-CENTRIC SECURITIZATION: THE DOMINANT DISCOURSE

In the contemporary geopolitical landscape, the securitization of AI is driven by a powerful coalition of state and corporate actors who frame the technology as the ultimate arbiter of global power. This dominant discourse prioritizes the “referent object” of the State (or the “Free World”) and views AI primarily as a strategic asset. A specific and highly influential strain of securitization has emerged from Silicon Valley and associated “AI Safety” institutes. This is the “Existential Risk” (X-Risk) narrative. Proponents, often linked to the ideologies of “Longtermism” and the “TESCREAL” bundle (Transhumanism, Extropianism, Singularitarianism, Cosmism, Rationalism, Effective Altruism, Longtermism), argue that future AI

systems could become “superintelligent” and misaligned with human values, leading to the extinction of the species ([Ord, 2020](#)).

This narrative constitutes a “Macro-Securitization” move, by framing the threat as total extinction, these actors successfully capture the attention of global policymakers, evident in events like the UK's AI Safety Summit ([Milmo & Stacey, 2023](#)). Critics, such as the Distributed AI Research Institute (DAIR), argue that this focus on hypothetical “God-like” AI serves as a distraction. It diverts political will and resources away from the actual harms being perpetrated by AI today (bias, labor exploitation) and centralizes power in the hands of the very tech companies building the systems, who position themselves as the only ones capable of “aligning” the beast they are creating ([DAIR, n.d.](#)). It effectively “securitizes the future” to “desecuritize the present.”

The most consequential state-level securitization is the framing of AI as the central theater of US-China rivalry. The US government views AI leadership as essential for maintaining the liberal international order. This has led to the “securitization of trade,” exemplified by the Biden administration's expansive export controls on advanced semiconductors (the “Chip War”) ([Glickman, 2024](#)). The argument is that high-end AI chips are inherently “dual-use” military goods, and thus their sale to China constitutes a national security threat. This narrative frames “US AI Leadership” as the referent object that must be protected at all economic costs ([Chander, 2025](#)).

The CCP (Chinese Communist Party) views AI through the lens of “Comprehensive National Security.” AI is seen as a tool for “Political Security”, maintaining the legitimacy of the Party through social governance (the Social Credit System) and information control ([Zeng, 2021](#)). The “AI Revolution in Military Affairs” is explicitly embraced by the People's Liberation Army (PLA) as a means to “leapfrog” US capabilities ([Glonek, 2024](#)). The Chinese state acts as the ultimate securitizing actor, mobilizing the entire society and economy to achieve AI supremacy as a matter of sovereignty ([Zeng, 2022](#)).

While less focused on the economic dimensions of AI, Russia has heavily securitized AI in the domain of Information Warfare. Russian intelligence agencies leverage AI for “astroturfing”, the automated creation of disinformation campaigns to sow discord in rival nations (Political

Security threat) ([Ünver, 2024](#)). They also utilize AI for predictive policing to preempt domestic public disturbances, securing the regime against internal dissent. The focus here is on asymmetric warfare, using AI to disrupt the democratic processes of adversaries ([Olteanu, 2024](#)).

6 THE HUMAN SECURITY IMPACT ASSESSMENT

If the state-centric view looks at AI from the “top down,” the human security view looks from the “bottom up.” By applying the seven pillars of the UNDP framework, it can map the granular, lived insecurities generated by the AI revolution.

Economic security aims to ensure a stable livelihood. AI is fundamentally reshaping the nature of work, creating new forms of “Algorithmic Insecurity.” Platforms such as Uber, Amazon Flex, and food delivery apps use “algorithmic management” to assign tasks, set wages, and monitor performance. These algorithms operate as “black boxes,” often terminating workers (“de-platforming”) for reasons that are opaque and unchallengeable ([Hickson & Morgana, 2025](#)).

This opaque control creates a psychological state of “algorithmic paranoia” among workers. They constantly try to guess what the algorithm wants, leading to hyper-vigilance and stress ([Alacovska et al., 2025](#)). The algorithm functions as a digital foreman that is invisible yet omnipresent.

Behind the perceived autonomy of AI systems lies a vast, hidden army of human laborers in the Global South who label data and moderate content for pennies. This “Ghost Work” is characterized by extreme precarity and exposure to traumatic content (filtering images of violence) ([Nothias, 2025](#)). The economic security of the Silicon Valley engineer is directly subsidized by the economic insecurity of the Kenyan data labeler.

Health security implies protection from disease and access to healthcare. AI is being integrated into diagnostic and resource allocation systems, often with devastating discriminatory effects. A landmark study revealed that an algorithm used by Optum to guide care for millions of Americans was systematically biased against patients with darker skin. The algorithm used “past healthcare costs” as a proxy for “future health needs.” Because the US

healthcare system historically underserves those patients with darker skin (spending less on them), the algorithm falsely concluded they were “healthier” than equally sick the other patients. As a result, patients with darker skin were denied access to high-risk care management programs ([Federspiel et al., 2023](#)). This is a clear case where AI actively degrades the health security of a specific racial group.

In conflict zones such as Palestine, the disparity in access to AI-driven healthcare technologies can exacerbate existing inequalities, creating a form of “medical apartheid” where one population benefits from precision medicine while another is excluded ([Gurupur et al., 2024](#)).

Social media algorithms, designed to maximize engagement (an economic metric), frequently promote content that harms mental health, such as “thinspiration” videos that fuel eating disorders in adolescents ([Costello et al., 2024](#)). Here, the commercial security of the platform is prioritized over the health security of the child.

In terms of food and environmental security, AI is often touted as a solution to climate change and hunger through “precision agriculture.” However, its own material footprint poses a threat. Training and running Large Language Models (LLMs) such as GPT-4 requires massive amounts of water to cool data centers. A conversation with a chatbot effectively “consumes” a bottle of water ([Drastig & Singh, 2025](#)). In water-scarce regions, data centers compete directly with local agriculture and communities for this vital resource, creating Environmental Insecurity ([Ward, 2023](#)). AI has improved famine prediction accuracy by 46% in pilot studies ([Balashankar et al., 2025](#)). However, prediction is not prevention. There is a risk that AI systems will be used to “triage” aid based on utilitarian calculations rather than human need, or that the “securitization of famine” will lead to military rather than humanitarian responses ([Ahmad et al., 2024](#)).

Regarding personal and community security, law enforcement agencies use AI to predict where crimes will occur. Because these systems are trained on historical arrest data, which is biased against poor and minority communities, they create a feedback loop. Police are sent to these neighborhoods, make more arrests, and feed that data back into the system ([Murray, 2010](#)). This securitizes the identity of the community itself,

turning neighborhoods into frameworks of algorithmic exclusion where residents are constantly watched but excluded from protection.

Future threat modeling suggests scenarios such as the “One Planeters,” a hypothetical wealthy fringe group using AI to attack space missions or other targets ([Johnson, 2019](#)). This illustrates how AI can empower non-state actors to disrupt community security.

In humanitarian crises, refugees (Rohingya, Syrians) are often forced to submit biometric data (iris scans) to receive food aid ([Paille *et al.*, 2024](#)). This data, ostensibly for “fraud prevention,” becomes a tool of surveillance. If shared with hostile home governments (as happened with the Rohingya and Myanmar), it poses a lethal threat to personal security ([Fejerskov *et al.*, 2021](#)).

Political security is related to the erosion of truth in this landscape. AI lowers the cost of generating high-quality disinformation (“deepfakes”). This threatens the “epistemic security” of democracies, the ability of citizens to know what is true ([Ünver, 2024](#)). Algorithmic curation creates “filter bubbles” that reinforce existing biases, deepening social polarization and making consensus impossible. The “gentle nudge” of digital affinity groups hardens into tribalism ([Sustainability Directory, 2025](#)).

7 DIGITAL COLONIALISM AND THE GEOPOLITICS OF DATA

A human security analysis must confront the global power dynamics embedded in AI. The relationship between the AI-rich Global North and the data-rich Global South is increasingly described as “Digital Colonialism” ([Nothias, 2025](#)). Digital colonialism functions through the extraction of data and the imposition of technological dependency. Just as colonial powers extracted rubber and gold, AI companies extract the “raw material” of data from the Global South to train their models, which are then sold back to those countries as finished products ([Nothias, 2025](#)).

Several cases should be mentioned at this point. Daniel Motaung, a South African content moderator for Facebook (via a subcontractor in Kenya), attempted to unionize to demand better mental health support for the trauma

of viewing graphic content. He was fired and blacklisted. His case illustrates the human cost of the “clean” AI interface ([Nothias, 2025](#)). This is the “hidden abode” of AI production. The Global South is becoming dependent on Northern cloud infrastructure and AI models. This lack of “Digital Sovereignty” means that the political and economic security of these nations is held hostage by foreign corporations ([de Freitas, 2025](#)).

In Kenya, the government's push for a biometric National ID (Huduma Namba) clashed with the UNHCR's biometric database of refugees. Many Kenyan citizens near the border, who had previously registered as refugees to access aid during droughts, found themselves blocked from getting National IDs because their biometrics were already in the “refugee” system ([Weitzberg, 2025](#)).

This “Double Registration” problem illustrates the rigidity of algorithmic categorization. A database error or a past survival strategy renders a citizen “stateless” in the eyes of the algorithm. Here, the pursuit of “state security” (secure borders, identification) directly destroys the “political security” (citizenship) of the marginalized.

8 NORMATIVE FUTURES: GOVERNING THE ALGORITHM

How do the global society govern these risks? Emerging regulatory frameworks offer competing visions of the future. The European Union has taken the lead with the AI Act, the first comprehensive AI law ([Hoffmann, 2024](#)). It adopts a risk-based approach, banning “unacceptable” risks (such as social scoring and real-time biometric identification in public spaces) and strictly regulating “high-risk” applications (such as migration and employment AI) ([Powell, 2024](#)). While it protects “Fundamental Rights,” it contains significant loopholes. “National Security” exemptions remain, and the protection of migrants is weaker than that of citizens ([Matlin et al., 2025](#)). It also allows for the export of banned technologies (such as surveillance tools) to non-EU countries, effectively exporting insecurity to the Global South ([Hoffmann, 2024](#)).

The African Union's AI Strategy emphasizes “Developmental Securitization.” It views AI primarily as a tool for achieving the Sustainable

Development Goals (SDGs) ([African Union Commission, 2024](#)). The strategy focuses on building “human capacity,” ensuring “data sovereignty” (keeping African data in Africa), and establishing bodies such as the African Observatory on Responsible AI ([Abiero *et al.*, 2024](#)). It seeks to prevent the continent from becoming a mere consumer of Western AI.

The UNESCO Recommendation on the Ethics of AI (2021) provides a global normative compass ([UNESCO, 2021](#)). Unlike the EU's market regulation, it focuses on universal values, human dignity, environmental sustainability, and gender equality. It explicitly links AI to the “protection of the environment” and calls for bans on AI for mass surveillance and social scoring ([UNESCO, 2021](#)). It serves as a crucial tool for nations developing their own regulations, offering a “human rights-based approach” that transcends the US-China binary.

9 CONCLUSION: TOWARD A HUMAN-CENTRIC SECURITY ARCHITECTURE

This chapter has argued that the securitization of artificial intelligence is a political process that masks a profound vertical security dilemma. By tracing the genealogy of computational control, from Cold War cybernetics to post-9/11 surveillance architectures, we have demonstrated how state fortification consistently breeds individual vulnerability. Central to this tension is the phenomenon of digital colonialism, where the extraction of data from the Global South serves to consolidate the power of Northern technological centers while exporting systemic insecurity to the margins.

Addressing these systemic violences requires moving beyond abstract existential narratives toward a pragmatic, human-centric security architecture. This demands the institutionalization of transparency in black-box algorithms, the enforcement of data sovereignty against colonial extractions, and the integration of rigorous human security audits into every stage of AI deployment, from precision agriculture to autonomous systems. Future governance must transcend national security exemptions to prioritize the safety, dignity, and agency of the individual over geopolitical competition.

The securitization of artificial intelligence, as this chapter has demonstrated, is not a neutral technical imperative but a profoundly political contest over who defines threat, who bears risk, and who reaps the rewards of technological progress. Through the integrative lens of securitization theory and the human security paradigm, we have traced the historical and discursive pathways that have elevated AI from a Cold War cybernetic experiment to a cornerstone of contemporary great power rivalry and everyday governance. From the offset strategies of the 1970s to the post-9/11 architectures of total information awareness, and from Project Cybersyn's foreclosed democratic potential to today's arms-race rhetoric, AI's evolution reveals a persistent pattern, technologies forged in the name of state survival are repurposed as instruments of control, often at the expense of individual dignity and agency.

The dominant state-centric and existential risk narratives, championed by geopolitical actors in the US-China theater, Russian information warriors, and Silicon Valley's longtermist vanguard, frame AI as a macro-threat to sovereignty or civilization itself. Yet, as our analysis of the seven pillars of human security underscores, this gaze obscures the tangible, immediate violences embedded in algorithmic deployment. Economic precarity stalks the gig worker ensnared by opaque platforms and the data laborer in the Global South, whose “ghost work” subsidizes Northern innovation. Health disparities widen as biased diagnostics ration care along racial lines, while environmental footprints of data centers parch communities in water-scarce regions. Personal and community securities fracture under predictive policing's feedback loops and biometric enclosures that render refugees doubly precarious. Political securities erode amid deepfake disinformation and filter-bubble tribalism, as digital colonialism extracts data sovereignty from the margins to fortify the centers of power.

This dual-use dilemma, wherein systemic fortification breeds subjective vulnerability, demands a radical normative reorientation. Desecuritizing routine harms such as algorithmic bias and labor extraction from the realm of “ethics” into urgent political imperatives, while securitizing data rights and epistemic injustices, offers a pathway to epistemic justice. The emerging regulatory mosaic, from the EU AI Act's risk tiers to the African Union's developmental safeguards and UNESCO's ethical universals, hints at plural futures. Yet, true governance must transcend exemptions for

“national security” and exportable insecurities, fostering global standards that prioritize human-centered metrics over competitive edges.

Securing AI requires elevating the human as the referent object of protection, not the state or the singularity. By democratizing transparency in black-box algorithms, enforcing data sovereignty against colonial extractions, and embedding human security audits into every deployment, from precision agriculture to lethal autonomous systems, the global society can transform AI from a vector of fragility into a scaffold for “freedom from fear and want.” The choice is not technological inevitability but collective agency, will the global society allow the algorithm to sort us into frameworks of algorithmic inclusion and exclusion, or harness it to weave a tapestry of shared security? The stakes are not merely existential but existentializing, shaping not just survival but the quality of life in a world where code writes our common fate.

The path toward a just digital order lies in the desecuritization of routine algorithmic harms, returning issues of labor rights, health equity, and bias to the realm of normal democratic deliberation. Future research should investigate the specific institutional mechanisms and global standards required to achieve such desecuritization in diverse political contexts. By elevating the human as the primary referent object of protection, global society can transform AI from a vector of fragility into a scaffold for shared dignity, ensuring that “freedom from fear and want” remains the guiding principle of the algorithmic age.

ACKNOWLEDGMENT

The authors used the Google Gemini 3 model solely for proofreading and readability enhancement of this manuscript. All conceptual development, arguments, interpretations, and conclusions are entirely the authors’ own, and full responsibility for the content remains with the authors.

REFERENCES

- Abiero, D., Karanja, N., and Kaaniru, J. (2024). *An in-depth analysis of the AU-AI continental strategy and implications on AI governance in the continent*. Centre for Intellectual Property and Information Technology Law. <https://cipit.strathmore.edu/an-in-depth-analysis-of-the-au-ai-continental-strategy-and-implications-on-ai-governance-in-the-continent/>
- Acheson, R. (2025). *WILPF inputs to the UN Secretary-General's report on artificial intelligence in the military domain*. Women's International League for Peace and Freedom. [https://docs-library.unoda.org/General Assembly First Committee - Eightieth session \(2025\)/79-239-WILPF-EN.pdf](https://docs-library.unoda.org/General_Assembly_First_Committee_-_Eightieth_session_(2025)/79-239-WILPF-EN.pdf)
- African Union Commission. (2024). *Continental artificial intelligence strategy: Harnessing AI for Africa's development and prosperity*. African Union. [https://au.int/sites/default/files/documents/44004-doc-EN- Continental AI Strategy July 2024.pdf](https://au.int/sites/default/files/documents/44004-doc-EN-Continental_AI_Strategy_July_2024.pdf)
- Ahmad, A., Liew, A.X.W., Venturini, F., Kalogeras, A., Candiani, A., Di Benedetto, G., Ajibola, S., Cartujo, P., Romero, P., Lykoudi, A., Mastrococco De Grandis, M., Xouris, C., Lo Bianco, R., Doddy, I., Elegbede, I., Falvo D'Urso Labate, G., García del Moral, L.F., and Martos, V. (2024). AI can empower agriculture for global food security: Challenges and prospects in developing nations. *Frontiers in Artificial Intelligence*, 7, Article 1328530. <https://doi.org/10.3389/frai.2024.1328530>
- Alacovska, A., Bucher, E., and Fieseler, C. (2025). Algorithmic paranoia: Gig workers' affective experience of abusive algorithmic management. *New Technology, Work and Employment*, 40, 421–435. <https://doi.org/10.1111/ntwe.12317>
- artificialintelligenceact.eu. (2024). *High-level summary of the AI Act*. <https://artificialintelligenceact.eu/high-level-summary/>
- Balashankar, A., Zimmer, P., Subramanian, L., and Fraiberger, S.P. (2025, June 12). Reading between the lines: Leveraging news data for AI-powered food insecurity forecasts. *VoxDev*. <https://voxdev.org/topic/methods-measurement/reading-between-lines-leveraging-news-data-ai-powered-food-insecurity>
- Buzan, B., Wæver, O., and de Wilde, J. (1998). *Security: A new framework for analysis*. Lynne Rienner Publishers.

- Chander, A. (2025, February 23). *The national security internet*. *Georgetown Law Journal*, 114. Georgetown University Law Center. <https://scholarship.law.georgetown.edu/cgi/viewcontent.cgi?article=3672&context=facpub>
- Costello, N., Sutton, R., Jones, M., Almassian, M., Raffoul, A., Ojumu, O., Salvia, M., Santoso, M., Kavanaugh, J.R., and Austin, S.B. (2024, February 12). Algorithms, addiction, and adolescent mental health: An interdisciplinary study to inform state-level policy action to protect youth from the dangers of social media. *American Journal of Law & Medicine*, 49(2–3), 135–172. <https://doi.org/10.1017/amj.2023.25>
- DAIR. (n.d.). *The TESCREAL bundle*. <https://www.dair-institute.org/projects/tescreal/>
- Daniels, O. (2022, December 21). *The AI “revolution in military affairs”*: What would it really look like? Center for Security and Emerging Technology. <https://cset.georgetown.edu/article/the-ai-revolution-in-military-affairs-what-would-it-really-look-like/>
- de Freitas, M.V. (2025, October 3). *Digital sovereignty and data colonialism: Shaping a just digital order for the Global South* [Policy paper]. Policy Center for the New South. <https://www.policycenter.ma/publications/digital-sovereignty-and-data-colonialism-shaping-just-digital-order-global-south>
- Drastig, K., and Singh, R. (2025). Review of water use assessment in livestock production systems and supply chains. *Water*, 17(19), Article 2819. <https://doi.org/10.3390/w17192819>
- Elahi, H., Osborne, M., and Thompson, J. (2022). *Total information awareness* [Art installation]. Grizzly Grizzly Annex. <https://roboticart.org/project/total-information-awareness/>
- Federspiel, F., Mitchell, R., Asokan, A., Umana, C., and McCoy, D. (2023). Threats by artificial intelligence to human health and human existence. *BMJ Global Health*, 8(5), Article e010435. <https://doi.org/10.1136/bmjgh-2022-010435>
- Fejerskov, A.M., Clausen, M.-L., and Seddig, S. (2021, August 17). *Risks of technology use in humanitarian settings: Avoiding harm, delivering impact* [Policy brief]. Danish Institute for International Studies. <https://www.diis.dk/en/research/risks-of-technology-use-in-humanitarian-settings-avoiding-harm-delivering-impact>

- Floyd, R. (2021). Securitisation and the function of functional actors. *Critical Studies on Security*, 9(1), 81–97. <https://doi.org/10.1080/21624887.2020.1827590>
- Glickman, S. (2024). *AI and tech industrial policy: From post-Cold War post-industrialism to post-neoliberal re-industrialization*. AI Now Institute. <https://ainowinstitute.org/publications/ai-and-tech-industrial-policy-from-post-cold-war-post-industrialism-to-post-neoliberal-re-industrialization>
- Glonek, J. (2024). The coming military AI revolution. *Military Review*. Army University Press. <https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2024/MJ-24-Glonek/>
- Gurupur, V., Cartwright, A., Johnson, M.D., Neubert, C., Oetjen, D.M., Robertson, M., and Nasternak, J. (2024, June). *UCF B.S. in health informatics degree program proposal* [Proposal]. University of Central Florida Board of Trustees. <https://bot.ucf.edu/wp-content/uploads/sites/5/2024/06/UCF-B.S.-in-Health-Informatics-Degree-Program-Proposal.pdf>
- Herrera, L. (2022, December 15). *Insights into Project Cybersyn* (Working Paper No. 67). Portland State University. https://pdxscholar.library.pdx.edu/econ_workingpapers/67
- Hickson, J., and Morgana, S. (2025). The global politics of precarity and insecure work: Introduction. *Global Political Economy*. Advance online publication. <https://doi.org/10.1332/26352257Y2025D000000042>
- Hoffmann, M. (2024, August 1). *The finalized EU artificial intelligence act: Implications and insights*. Center for Security and Emerging Technology. <https://cset.georgetown.edu/article/the-finalized-eu-artificial-intelligence-act-implications-and-insights/>
- Johnson, B.D. (2019). *Information disorder machines: Weaponizing narrative and the future of the United States of America*. Arizona State University Threatcasting Lab. <https://threatcasting.asu.edu/sites/g/files/litvpz1036/files/2019-11/01threatcasting-2019-IDM.pdf>
- Joseph, P. (2023, July 7). The story of Stafford Beer's “Project Cybersyn.” *Medium*. <https://peter-joseph.medium.com/the-story-of-stafford-beers-project-cybersyn-4bdbfabd78eb>

- Karyotis, G., and Paterson, I. (2025). Understanding securitization success: A new analytical framework. *International Studies Review*, 27(1), Article viaf006. <https://doi.org/10.1093/isr/viaf006>
- Koundouri, P., Aslanidis, P.-S., Dellis, K., Plataniotis, A., and Feretzakis, G. (2025). Mapping human security strategies to sustainable development goals: A machine learning approach. *Discover Sustainability*, 6(96). <https://doi.org/10.1007/s43621-025-00883-w>
- Lupovici, A. (2021). The dual-use security dilemma and the social construction of insecurity. *Contemporary Security Policy*, 42(3), 257–285. <https://doi.org/10.1080/13523260.2020.1866845>
- Matlin, S.A., Claron, I.M.M., Merone, J., Netto, G., Takian, A., Zaman, M.H., and Saso, L. (2025, August 8). Artificial intelligence in migrant health: A critical perspective on opportunities and risks. *The Lancet Regional Health – Europe*, 57, Article 101421. <https://doi.org/10.1016/j.lanepe.2025.101421>
- Milmo, D., and Stacey, K. (2023, September 25). *Experts disagree over threat posed but artificial intelligence cannot be ignored*. *The Guardian*. <https://www.theguardian.com/technology/2023/sep/25/experts-disagree-over-threat-posed-but-artificial-intelligence-cannot-be-ignored-ai>
- Morozov, E., and Higgins, M. (2023, September 29). Why Chile's Santiago Boys launched Project Cybersyn. *The World Today*. Chatham House. <https://www.chathamhouse.org/publications/the-world-today/2023-10/why-chiles-santiago-boys-launched-project-cybersyn>
- Murray, N. (2010). Profiling in the age of total information awareness. *Race & Class*, 52(2), 3–19. <https://doi.org/10.1177/0306396810377002>
- Naghbeyli, T. (2025, November 28). *In which ways could Copenhagen School's analysis of security be critiqued by alternative proponents of a new security paradigm?* Politicon.co. <https://politicon.co/en/essays/64/in-which-ways-could-copenhagen-schools-analysis-of-security-be-critiqued-by-alternative-proponents-of-a-new-security-paradigm>
- Nothias, T. (2025). An intellectual history of digital colonialism. *Journal of Communication*. Advance online publication. <https://doi.org/10.1093/joc/jqaf003>
- Olteanu, M. (2024, December 30). The Copenhagen School's widening security theory in relation to cybersecurity. Applicability and

- implications. *Security and Defence Quarterly*, 48(4).
<https://doi.org/10.35467/sdq/193049>
- Ord, T. (2020). *The precipice: Existential risk and the future of humanity*. Hachette Books.
- Paille, P., Toole, H., Politi, C., and Ohrvik-Stott, J. (2024). *Humanitarian technology adoption case study: Biometrics* (Research Report RRA3192-6). RAND Corporation.
https://www.rand.org/content/dam/rand/pubs/research_reports/RRA3100/RRA3192-6/RAND_RRA3192-6.pdf
- Powell, R. (2024, August). *The EU AI Act: National security implications*. Centre for Emerging Technology and Security.
<https://cetas.turing.ac.uk/publications/eu-ai-act-national-security-implications>
- Rinaldi, A., and Teo, S.A. (2025, April 4). The use of artificial intelligence technologies in border and migration control and the subtle erosion of human rights. *International & Comparative Law Quarterly*, 74(1), 61–89. <https://doi.org/10.1017/S0020589325000090>
- Sustainability Directory. (2025). *What role does AI play in social polarization?* Lifestyle Sustainability Directory.
<https://lifestyle.sustainability-directory.com/question/what-role-does-ai-play-in-social-polarization/>
- Tucker, C.A. (2025, October 30). *Project Cybersyn: From 1970s economic cybernetics to 21st-century socio-technical innovation* [Conference paper]. ResearchGate.
https://www.researchgate.net/publication/397059976_Project_Cybersyn-From_1970s_Economic_Cybernetics_to_21st-Century_Socio-Technical_Innovation
- Ulloa Ruiz, M.A., and Bas Graells, G. (2025). The securitization of artificial intelligence: An analysis of its drivers and consequences. *Revista de Estudios Sociales*, 93.
<https://doi.org/10.7440/res93.2025.04>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*.
<https://unesdoc.unesco.org/ark:/48223/pf0000386510>
- United Nations Development Programme. (1994). *Human development report 1994*. Oxford University Press.
<https://hdr.undp.org/content/human-development-report-1994>

- Ünver, H.A. (2024, May). *Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights* (Study No. PE 754.450). European Parliament, Directorate-General for External Policies. <https://doi.org/10.2861/907329>
- Ward, F.A. (2023). Innovations for the water resource economics curriculum: Training the next generation. *Applied Economics Teaching Resources*, 5(1), 38–52. https://www.aetrjournal.org/UserFiles/file/AETR_2022_038%20Full%20PDF%20Version.pdf
- Weitzberg, K. (2025, July 8). Keeping people out of camps: Biometric technologies, contested sovereignty, and border practices within humanitarian spaces. *Journal of Ethnic and Migration Studies*, 51(14), 3590–3609. <https://doi.org/10.1080/1369183X.2025.2513155>
- Wells, W. (2025). Battlefield evidence in the age of artificial intelligence-enabled warfare. *Chicago Journal of International Law*, 26(1), 249–280. <https://cjl.uchicago.edu/print-archive/battlefield-evidence-age-artificial-intelligence-enabled-warfare>
- Zeng, J. (2021). Securitization of artificial intelligence in China. *The Chinese Journal of International Politics*, 14(3), 417–445. <https://doi.org/10.1093/cjip/poab005>
- Zeng, J. (2022). *Artificial intelligence with Chinese characteristics: National strategy, security and authoritarian governance*. Palgrave Macmillan. <https://doi.org/10.1007/978-981-19-0722-7>

Simulating Decision-Making at the Edge of Sovereignty Through Artificial Intelligence: Theoretical Narrative and Integration of Hyderabad

Abhishek Priyadarshi, Rachita Singh, Parakh Shrinag, Riddhi Bhardwaj,
Sahaj Thakur, Priyanjana Ghosh, Atharv K Arya, and Aditya Nair

University of Delhi

DOI: [10.1201/9781042018116-11](https://doi.org/10.1201/9781042018116-11)

ABSTRACT: This paper analyzes the dramatic struggle over the political future of the Hyderabad state in 1947–48, highlighting how Sardar Vallabhbhai Patel neutralized the Nizam's attempt to secure independence and explore alignment with Pakistan. Drawing on detailed records of negotiations, constitutional exchanges, and military planning, it situates Hyderabad's crisis within the broader uncertainty created by the lapse of British paramountcy. The Nizam's strategy rested on legal argumentation, international advocacy, and reliance on external actors, while the Government of India emphasized territorial integrity, popular security, and the imperative of responsible government. Hyderabad's anomalous status—a vast, landlocked princely state with a Muslim ruling elite governing an overwhelmingly Hindu population—rendered its secessionist posture especially destabilizing for the nascent Indian Union. The paper further places the Hyderabad confrontation in the wider strategic environment of the partition-era conflicts, particularly the war in Jammu and Kashmir. It engages with accounts suggesting that Patel maintained Indian troops in readiness in Srinagar to deter Pakistan, even as the Nizam mobilized state forces and Razakar militias. This simultaneity reveals Patel's calibrated statecraft: sustained negotiations and standstill arrangements were combined with credible military preparedness to signal resolve. When diplomacy collapsed amid escalating internal violence, the Indian leadership authorized ‘police action’ (Operation Polo), swiftly extinguishing the secessionist challenge. The Hyderabad episode thus illuminates Patel's fusion of legalism, political resolve, and coercive

capacity in consolidating India's sovereignty at a moment of profound constitutional and territorial fragility.

Keywords: Hyderabad State, Sardar Vallabhbhai Patel, Indian integration, Operation Polo, Partition-era security

1 METHODOLOGY

This study employs a qualitative historical methodology based on systematic archival research of the Ministry of States, Hyderabad Branch files at the National Archives of India, including constitutional, diplomatic, financial, and administrative records relating to the Hyderabad crisis. To address the volume and heterogeneity of these sources, AI-assisted analytical tools are used as a heuristic aid for text parsing and thematic clustering, enabling the identification of recurring patterns of negotiation, escalation, and institutional breakdown over time. AI is not treated as a historical actor but as a methodological extension of conventional archival practice, enhancing analytical clarity by revealing decision thresholds and structural regularities that are often obscured in linear narrative accounts, while remaining firmly grounded in contextual interpretation and empirical evidence.

2 INTRODUCTION

This article employs Artificial Intelligence-inspired decision modelling as an interpretive framework rather than a historical claim. By mapping historical actors onto functional categories familiar in contemporary AI systems—such as rule-based agents, constraint-sensitive optimizers, and execution modules—the analysis clarifies how authority, information flow, and escalation thresholds operated during the Hyderabad crisis. This approach does not anachronistically attribute artificial intelligence to historical actors; instead, it uses AI as an analytical vocabulary to expose structural features of decision-making that conventional narrative history often leaves implicit. The Hyderabad case, in this sense, offers an early illustration of human-centred governance architectures later formalized in computational systems.

Table 1. Actors, Decision Logics, and Governance Functions in the Integration of Hyderabad (Interpreted through an AI-Analytical Lens).

Personality	Historical Role	Decision Logic (Human)	AI-Analytical Interpretation	Governance Insight
Sardar Vallabhbhai Patel	Final authority; authorized Operation Polo	Threshold-based escalation; risk acceptance	Rule-based agent with override authority once system instability exceeds tolerance	Sovereignty enforcement requires a terminal decision node
V. P. Menon	Chief negotiator and strategist	Iterative bargaining; scenario evaluation	Sequential decision optimizer using feedback loops	Bureaucratic intelligence precedes coercion
Jawaharlal Nehru	PM; international legitimacy	Constraint-sensitive deliberation	Ethics-constrained policy model	Liberal norms delay but do not prevent consolidation
Lord Mountbatten	Mediator	Conflict dampening; delay optimization	Human-in-the-loop moderator	Mediation fails when objectives are non-convergent
Nizam of Hyderabad	Claimant of sovereignty	Legal formalism; externalization	Model locked in legacy parameters	Static rule systems fail in dynamic environments
Sir Walter Monckton	Legal architect	Textual–treaty reasoning	Symbolic reasoning engine	Law without enforcement capacity is brittle

Personality	Historical Role	Decision Logic (Human)	AI-Analytical Interpretation	Governance Insight
Nawab of Chhatari	Moderate constitutionalist	Compromise-seeking	Low-weight decision node	Moderation collapses without control over force
Kasim Razvi	Militant mobilizer	Ideological absolutism	Adversarial actor amplifying system noise	Extremism accelerates system breakdown
Mir Laik Ali	PM (later phase)	Internationalization; delay	External arbitration subroutine	Forum shopping increases resolution cost
K. M. Munshi	Agent-General	Ground-truth reporting	Real-time anomaly detector	Field data overrides diplomatic assurances
Maj. Gen. J. N. Choudhuri	Military executor	Rapid force projection	Execution module	Coercion is effective when tightly bounded

Using AI-assisted analytical methods, this study constructs a chronological flow of Hyderabad's integration that goes beyond narrative description to offer a structured reconstruction of decision-making under stress. Artificial intelligence was employed as a heuristic and organizational tool to parse dense archival material—such as the failed negotiations of 11 July 1947, the rejection of treaty sovereignty on 12 August 1947, and the repeated collapse of draft “Heads of Agreement” between September 1947 and June 1948—and to identify recurring decision points across time. AI-assisted pattern recognition enabled the clustering of events into sequential phases: initial constitutional engagement (July–August 1947), internal radicalization marked by the rise of the Razakars and the resignation of moderate interlocutors on 25 August 1947, diplomatic exhaustion following repeated breaches of the Standstill Agreement after 29 November 1947, and finally the terminal phase of coercive resolution culminating in Operation Polo (13–17 September 1948). Rather than treating the

integration of Hyderabad as a singular act of force, this method reveals a rule-like progression in which constitutional mechanisms repeatedly failed under conditions of escalating internal violence, internationalization of the dispute through the United Nations in August 1948, and the erosion of responsible government, leading to Sardar Vallabhbhai Patel's authorization of military intervention on 9 September 1948. This approach does not anachronistically attribute artificial intelligence to historical actors; instead, it employs AI-inspired analytical modeling to illuminate the underlying decision logic, thresholds, and feedback loops that shaped human statecraft during the crisis, thereby enhancing interpretive clarity while remaining firmly grounded in archival evidence.

Sequential Collapse of Constitutional Negotiation and the Integration of Hyderabad (1947–48)



Figure 1. Sequential Collapse of Constitutional Negotiation and the Integration of Hyderabad (1947–48).

We applied AI in this study as a methodological aid. It enabled the structured extraction of information from heterogeneous archival files. These included constitutional correspondence, security-related memoranda, and records on property, trusts, and finance. AI-assisted clustering helped group recurring themes across these diverse sources. This process revealed patterns of negotiation, escalation, and institutional consolidation. Such patterns are not easily visible through linear narrative analysis alone.

The archival table below delineates the primary sources consulted in this study and specifies their analytical relevance. It consolidates heterogeneous records from the Ministry of States, Hyderabad Branch—including files on constitutional negotiations, financial arrangements, property and trust matters, foreign engagements, and post-accession settlements—and links them to the thematic categories employed in the analysis. By systematizing these materials, the table demonstrates the depth and range of archival engagement and clarifies how diverse documentary evidence informs the reconstructed chronology. This explicit mapping also underpins the AI-assisted methodology adopted in the study, indicating that thematic extraction and clustering were derived from identifiable primary sources rather than from abstract or speculative inference.

Table 2. Archival Sources from the Ministry of States, Hyderabad Branch (NAI) used for AI-Assisted Chronological and Analytical Reconstruction.

Sl. No.	Archival File	File Number	Thematic Focus	Use in the Article
1	Nizam Property Files	F 57-H/49	Trusts, private property, family settlements	Establishes post-integration negotiations over sovereignty, legitimacy, and property rights
2	Consent of the Nizam for Donation to Banaras Hindu University	F.4(20)-H/1949	Philanthropy and cultural patronage	Demonstrates continued negotiation of authority and symbolic legitimacy after integration

Sl. No.	Archival File	File Number	Thematic Focus	Use in the Article
3	Prince of the Bera Trust	F.1(58)-H/1949	Trust creation and asset management	Used to analyse fragmentation of sovereign authority into legal-administrative forms
4	Acquisition of Immovable Property by the Nizam (Bombay)	F.10(56)/1949	Property acquisition outside Hyderabad	Illustrates limits placed on princely autonomy in post-accession India
5	Question Arising out of the Creation of Trust by the Nizam	(60)-H/1949	Legal disputes over trusts	Shows administrative consolidation replacing dynastic discretion
6	Provisional Accession of Hyderabad & Draft Constitution for States	(64)-H/1949	Constitutional transition	Core evidence for the phased constitutional integration model
7	Declaration of Foreign Holdings Abroad	7.1(133)-H/1949	External financial interests	Supports the argument on externalisation of sovereignty and security concerns
8	Programme of the Egyptian Journalist Delegation Visiting India	12(12)-H/1949	Foreign media and diplomacy	Used to demonstrate international perception management during integration
9	Grant Made by the Nizam to Jama Masjid, Delhi	2(1)-H/1949	Religious and political patronage	Indicates negotiation of minority legitimacy within the Indian Union

Sl. No.	Archival File	File Number	Thematic Focus	Use in the Article
10	Hyderabad: ₹20 Crore Loan to Pakistan	2(2)-H/1949	Financial–strategic alignment	Critical evidence of security escalation and diplomatic rupture
11	Arrangement & Fixation of Privy Purse under Article 291	2(1)-H/1950	Financial settlement & sovereignty	Demonstrates institutionalization of accession and closure of dynastic claims
12	Acceptance of Draft Constitution by the Nizam	1(64)-Hyderabad/1949	Formal constitutional consent	Marks completion of constitutional integration phase

3 HISTORICAL ORIGINS, POLITICAL FRAGILITY, AND THE CRISIS OF ACCESSION

The princely state of Hyderabad emerged from the gradual unravelling of Mughal imperial authority in the Deccan, a process shaped by military entrepreneurship, negotiated sovereignty, and the growing intervention of European powers. Its origins can be traced to the career of Ghaziud-din Khan Feroz Jang, a formidable general under Emperor Aurangzeb, who claimed descent from Abu Bakr, the first Khalifa—an ancestry that later assumed political and symbolic importance in legitimizing Asaf Jahi rule. In 1713, six years after Aurangzeb's death, Emperor Farrukhsiyar appointed Ghaziud-din as Viceroy of the Deccan with the title *Nizam-ul-Mulk* (Regulator of the Realm), while Emperor Muhammad Shah later bestowed upon him the honorific *Asaf Jah* ([Richards, 1993](#)). By 1724, Ghaziud-din's son, Mir Qamruddin Chin Qilich Khan, had effectively severed Mughal control and established Hyderabad as an independent state. Yet, in a characteristic display of early modern political pragmatism, the Asaf Jahi rulers continued to profess nominal allegiance to the Mughal throne until 1858, long after real authority had evaporated ([Dalrymple, 2019](#)).

This duality—formal autonomy coupled with symbolic subordination—was a defining feature of Hyderabad's political evolution. Despite ruling for over two centuries, the Asaf Jahi dynasty never fully consolidated sovereignty in the manner of centralized territorial states. Succession disputes were frequent, and internal legitimacy remained fragile. To survive these challenges, the Nizams consistently

relied on external powers: initially the French, later the British, and at critical junctures, regional military actors such as Hyder Ali's Mysore ([Sarkar, 1983](#)). While such alliances preserved the dynasty, they hollowed out its autonomy. Over time, Hyderabad became less a sovereign power and more a strategically useful intermediary polity, whose survival depended on external guarantees rather than internal cohesion.

British paramountcy institutionalized this dependence. The Nizam was transformed from a semi-independent ruler into a protected ally within the imperial order. This transformation was most clearly expressed in the 1902 Berar agreement under Lord Curzon, which reaffirmed the Nizam's sovereignty in theory while leasing the strategically and economically vital province of Berar to the Government of India in perpetuity. In return for an annual rent, the British assumed full administrative control, and the Hyderabad Contingent—once a symbol of princely military autonomy—was effectively absorbed into the Indian Army, with the Government of India assuming responsibility for the defense of the Nizam's dominions ([Menon, 1956](#)). By the early twentieth century, Hyderabad's sovereignty existed more on paper than in practice.

Nevertheless, Hyderabad remained an imposing entity. By the 1940s it was the largest princely state in peninsular India and the second largest overall after Jammu and Kashmir. Its landlocked geography, surrounded entirely by British Indian provinces—Bombay, Madras, and the Central Provinces—placed it at the strategic heart of the subcontinent. According to the 1941 census, Hyderabad had a population of approximately 16.34 million and an annual revenue of about Rs. 26 crores. It possessed its own currency, paper money, and postal system, reinforcing the Nizam's belief in his state's distinct international personality. Yet beneath these trappings of sovereignty lay a profound structural contradiction: while over 85 percent of the population was Hindu, the civil services, police, and armed forces were overwhelmingly Muslim ([Copland, 1988](#)). This imbalance would become politically explosive during the transition from empire to nation-state.

The lapse of British paramountcy in 1947 exposed these accumulated contradictions with unprecedented force. Following the British government's announcement of 3 June 1947, the seventh Nizam, Mir Usman Ali Khan, issued a *firman* declaring that Hyderabad would neither join India nor Pakistan and would instead become independent on 15 August 1947. This declaration reflected not merely dynastic ambition but a profound misreading of the postcolonial order. The Nizam's subsequent protest against Clause 7 of the Indian Independence Bill—which explicitly denied dominion status to princely states—revealed a reliance on legal formalism divorced from political realities ([Menon, 1956](#)). His decision to hire Sir Walter Monckton, one of Britain's most eminent constitutional lawyers, further underscored this approach. The Nizam sought salvation through legal argument and

imperial-era intermediaries at a moment when power was rapidly shifting to mass-based, territorially consolidated states.

For the leadership of the Indian Union, Hyderabad's bid for independence was intolerable. An autonomous Hyderabad would have geographically bisected India, disrupted communications between northern and southern India, and set a dangerous precedent for other princely states. Sardar Vallabhbhai Patel famously described such an outcome as a “cancer in the belly of India.” Yet Patel, along with Prime Minister Jawaharlal Nehru, initially pursued a strategy of prolonged negotiation, consistent with their commitment to constitutionalism, democratic legitimacy, and international credibility ([Guha, 2007](#)). These negotiations—standstill agreements, draft treaties, and repeated concessions—were not mere delaying tactics but genuine attempts to integrate Hyderabad peacefully.

However, diplomacy unfolded against a deteriorating internal situation. The rise of the Ittehad-ul-Muslimeen and its paramilitary wing, the Razakars, introduced an overtly communal and coercive element into Hyderabad's politics. Their militant posture, claims of Islamic sovereignty, and intimidation of the Hindu majority undermined the Nizam's assertion that he could maintain law and order. Rather than strengthening Hyderabad's bargaining position, the Razakars delegitimized it in the eyes of both the Indian leadership and the international community.

The Hyderabad crisis thus exposes the deeper paradox of princely sovereignty at the moment of decolonization. Hyderabad claimed historical legitimacy and legal autonomy but depended on coercive minorities, external patrons, and colonial-era protections for survival. Its failure was not primarily the result of Indian coercion but of an inability to reconcile dynastic privilege with the imperatives of popular governance, demographic reality, and territorial consolidation. The state's eventual absorption into the Indian Union was therefore not merely an act of force but the culmination of a long historical trajectory in which political fragility, structural imbalance, and misplaced legalism rendered independence unsustainable.

4 EARLY NEGOTIATIONS AND DIPLOMATIC DEADLOCK OVER HYDERABAD

The earliest phase of the Hyderabad crisis unfolded through a series of formal diplomatic engagements in July and August 1947, revealing the inherent limits of constitutional negotiation when confronted with fundamentally incompatible political objectives. Far from being a simple disagreement over modalities of accession, these exchanges exposed a deeper clash between a dynastic conception of sovereignty and the imperatives of territorial consolidation underpinning the postcolonial Indian state. Archival records from the Ministry of States indicate that the first substantive

engagement took place on 11 July 1947, when a Hyderabad delegation led by the Nawab of Chhatari, President of the Executive Council, met representatives of the Government of India in New Delhi (Government of India, Ministry of States, *Hyderabad Negotiations*, Note of Meeting, 11 July 1947). The delegation—comprising Sir Walter Monckton, K. C. Abdur Rahim, Nawab Ali Yawar Jung, and Pingle Venkatarama Reddy—declined to commit to accession, and the official record tersely notes that the discussions “reached nowhere” ((Government of India, Ministry of States, *Hyderabad Negotiations*, Note of Meeting, 11 July 1947). This laconic phrasing itself reflects the depth of the impasse rather than a lack of engagement.

During these talks, the Hyderabad delegation adopted a strategy of calculated ambiguity. While stopping short of an outright rejection of accession to India, it suggested that excessive pressure from New Delhi might compel the Nizam to consider accession to Pakistan. This assertion was less a realistic policy option than a bargaining tactic intended to leverage Pakistan's existence as a countervailing force. Lord Mountbatten's response, preserved in the same archival note, was both legalistic and strategically grounded. He acknowledged that the Nizam was technically entitled to accede to Pakistan under the prevailing constitutional framework, but emphasized the “mechanical difficulties presented by geography”—a diplomatic formulation that pointed to Hyderabad's landlocked position in the Indian heartland and the practical impossibility of sustaining such an arrangement (*ibid.*). This exchange marked an early recognition by British authorities that the formal legal equality of princely states could not override the geopolitical logic of territorial contiguity and administrative viability.

Efforts to normalize Hyderabad's position within the wider process of princely integration followed shortly thereafter. On 25 July 1947, V. P. Menon proposed that the Nawab of Chhatari be included in a Negotiation Committee alongside representatives of other princely states, an initiative designed to embed Hyderabad within a collective framework and dilute its claim to exceptional treatment (Menon, Memorandum to Nawab of Chhatari, 25 July 1947, Ministry of States Papers). The Nawab's refusal to participate was politically significant. It underscored the Nizam's determination to preserve a bilateral negotiating posture and avoid any mechanism that might implicitly subordinate Hyderabad's claims to those of other rulers. This refusal reinforced Indian perceptions that Hyderabad was seeking to remain outside the emerging constitutional consensus rather than merely negotiating the terms of entry (Menon, Memorandum to Nawab of Chhatari, 25 July 1947, Ministry of States Papers).

As the transfer of power approached, the depth of the impasse became increasingly unmistakable. In a letter dated 8 August 1947, the Nizam wrote directly to Lord Mountbatten stating that he could not contemplate bringing Hyderabad into an “organic union” with India, though he expressed willingness to enter into a treaty-

based relationship (Nizam of Hyderabad to Mountbatten, 8 August 1947, Ministry of States Papers). The language of the letter revealed a continued reliance on pre-independence notions of indirect rule and treaty sovereignty. Mountbatten rejected this proposal on 12 August 1947, making clear that treaty sovereignty for a princely state was incompatible with the constitutional structure of the impending dominion and with the principle that sovereignty would henceforth reside in territorially consolidated states rather than dynastic entities (Mountbatten to Nizam, 12 August 1947, Ministry of States Papers).

On 15 August 1947, the day of Indian independence, Lord Mountbatten formally acknowledged the failure to resolve the Hyderabad issue within the colonial framework. He requested a two-month extension from the Government of India to continue negotiations, a request approved by the Indian cabinet in the interest of exhausting all peaceful and constitutional options (Cabinet Secretariat Note, 15 August 1947, Government of India). On the same day, Mountbatten addressed the Constituent Assembly and publicly stated that, with the sole exception of Hyderabad, all princely states had acceded to India—a declaration that carried both symbolic and political weight (Constituent Assembly Debates, 15 August 1947). Hyderabad was thus publicly identified as an anomaly within the new constitutional order, even as negotiations formally continued.

Despite this growing isolation, Hyderabad made a final attempt to revive diplomatic engagement. On 17 August 1947, the Nawab of Chhatari again wrote to Mountbatten expressing a desire to resume negotiations, and Mountbatten agreed to receive the delegation on 25 August 1947 (Nawab of Chhatari to Mountbatten, 17 August 1947; Mountbatten's reply, Ministry of States Papers). This exchange illustrates the oscillatory nature of Hyderabad's strategy—alternating between defiance and conciliation—without fundamentally altering its core position on sovereignty.

By late August 1947, Indian policymakers had largely concluded that Hyderabad's leadership was unwilling, or unable, to reconcile dynastic sovereignty with the imperatives of territorial unity, popular legitimacy, and democratic governance. The failure of these early negotiations did not merely reflect diplomatic exhaustion; it marked the structural collapse of a constitutional approach predicated on exceptionalism. These exchanges laid the groundwork for the subsequent Standstill Agreement and, ultimately, for the conclusion within the Indian leadership that coercive integration might prove unavoidable.

5 THE UNFOLDING OF THE HYDERABAD CRISIS: FROM CONSTITUTIONAL ENGAGEMENT TO STRATEGIC

BREAKDOWN

The Hyderabad crisis did not unfold as a sudden breakdown of relations between the princely state and the Government of India; rather, it evolved through a cumulative erosion of constitutional negotiation under the pressures of internal radicalization, strategic mistrust, and irreconcilable conceptions of sovereignty. By late August 1947, the limits of legalistic diplomacy had become increasingly apparent, as the Nizam's insistence on exceptional status collided with the Indian leadership's determination to consolidate territorial integrity in the aftermath of Partition.

A decisive moment in this unfolding crisis occurred on 25 August 1947, when Sir Walter Monckton, the Nawab of Chhatari, and Nawab Ali Yawar Jung resigned from their official positions. Archival records indicate that these resignations followed orchestrated attacks in the Hyderabad press led by the Ittehad-ul-Muslimeen, which targeted moderate officials advocating accommodation with India (Government of India, Ministry of States, 1947). This episode marked a critical shift in the internal political balance of Hyderabad, signaling the growing ascendancy of extremist elements and the corresponding marginalization of constitutional intermediaries. The Nizam's subsequent appeal to Lord Mountbatten to persuade Monckton to remain in service—and Monckton's conditional agreement following a formal condemnation of the attacks—demonstrated both the fragility of Hyderabad's negotiating apparatus and its dependence on external mediation (Government of India, Ministry of States, 1947).

The internal deterioration of governance in Hyderabad was inseparable from the rise of the Ittehad-ul-Muslimeen under the leadership of Kasim Razvi. Archival descriptions depict Razvi as combining ideological absolutism with political opportunism, mobilizing mass support through militant rhetoric and the creation of the Razakar militia (Government of India, Ministry of States, 1947). The routine militarization of public life and the explicit advocacy of a theocratic political order fundamentally undermined the Nizam's claims to maintaining law and order. For Indian policymakers, these developments transformed Hyderabad from a constitutional anomaly into a potential security liability within the heart of the subcontinent.

In an effort to arrest this downward spiral, Lord Mountbatten and V. P. Menon travelled to Hyderabad on 26 August 1947. These discussions, however, failed to produce any substantive convergence. Archival notes emphasize that Sardar Vallabhbhai Patel remained resolute in his position that only full accession could provide a viable constitutional solution (Government of India, Ministry of States, 1947). This insistence reflected a broader recalibration within the Indian leadership: patience with princely exceptionalism was increasingly subordinated to concerns over internal stability and national cohesion.

On 27 August 1947, Mountbatten proposed a referendum under British supervision as a potential mechanism to resolve the deadlock through popular consent. The Nizam's refusal of this proposal was revealing. It indicated not only a rejection of Indian authority but also a reluctance to submit princely sovereignty to democratic validation. In a state where the ruling elite was demographically and politically detached from the majority population, this refusal further eroded Hyderabad's moral legitimacy (Government of India, Ministry of States, 1947).

Negotiations continued into September but increasingly assumed a ritualistic character. On 8 September 1947, Monckton and the Nawab of Chhatari reiterated their preference for a treaty-based relationship with India. Mountbatten's response was unequivocal: Patel would accept nothing short of accession, and alternative formulations merely postponed an inevitable reckoning (Government of India, Ministry of States, 1947). During this meeting, Mountbatten also raised concerns regarding an arms order valued at approximately £3 million placed by Hyderabad with Czechoslovakia. Although the Nawab claimed that the order lacked final authorization, the episode reinforced Indian suspicions that Hyderabad was preparing for confrontation rather than accommodation (Government of India, Ministry of States, 1947).

The Nizam's correspondence during this period further accentuated these concerns. In a letter dated 18 September 1947, he warned that accession would precipitate widespread violence and sought Pakistan's intervention by requesting Muhammad Ali Jinnah to depute Sir Zafrullah Khan to Hyderabad. Jinnah's refusal—citing Khan's diplomatic obligations at the United Nations—highlighted both the limits of Pakistan's involvement and Hyderabad's growing diplomatic isolation (Government of India, Ministry of States, 1947).

Subsequent negotiations confirmed the exhaustion of constitutional possibilities. On 22 September 1947, Sir Sultan Ahmed advanced the argument that Indian pressure would provoke mass violence. Mountbatten's recorded response marked a significant moral inflection point: he questioned whether the Government of India could be expected to remain passive in the face of large-scale violence against the Hindu population of Hyderabad (Government of India, Ministry of States, 1947). The emphasis on civilian protection increasingly framed India's strategic calculus.

The final attempt at compromise came on 23 September 1947, when Monckton submitted another draft “Heads of Agreement.” Acting on Menon's advice, Mountbatten rejected the proposal the following day, concluding that incremental concessions merely prolonged deadlock without addressing the fundamental issue of sovereignty (Government of India, Ministry of States, 1947). On 26 September 1947, Monckton reported that the Nizam's position had hardened further and that Pakistan's influence had become “paramount,” effectively confirming Indian fears of external alignment.

By the end of September 1947, the Hyderabad issue had ceased to be a question of constitutional negotiation and had become a matter of state security. The crisis unfolded through a progressive collapse of legitimacy—internal, diplomatic, and moral—rather than through a single catalytic event. The failure of this phase laid the structural foundations for the Standstill Agreement and, ultimately, for coercive integration. In this sense, Hyderabad's trajectory illustrates the limits of dynastic sovereignty in an era defined by mass politics, territorial consolidation, and postcolonial state formation.

6 FROM STANDSTILL TO STRATEGIC EXHAUSTION: THE FINAL COLLAPSE OF NEGOTIATION

Following the collapse of negotiations in late September 1947, the Hyderabad question entered a prolonged and increasingly brittle phase marked by tactical manoeuvring, repeated breaches of assurance, and the cumulative erosion of trust between the Nizam's government and the Indian Union. What distinguished this phase was not the abandonment of negotiation but its degeneration into an instrument of delay. Dialogue continued, but it increasingly functioned as a means of postponing decision rather than resolving conflict, even as internal disorder within Hyderabad and external alignments steadily foreclosed the possibility of a peaceful constitutional settlement.

An early indication of this pattern emerged in October 1947, when V. P. Menon was invited to Hyderabad by the Nawab of Chhatari, only for the invitation to be abruptly cancelled on 6 October 1947 (Government of India, Ministry of States, 1947). This reversal reinforced Indian perceptions that the Hyderabad leadership was either unwilling or incapable of sustaining coherent engagement. The cancellation was not merely procedural; it reflected the growing fragmentation of authority within Hyderabad, where moderate constitutional voices were increasingly constrained by militant elements. When Lord Mountbatten met the Hyderabad delegation on 10 October 1947, the impasse persisted. Archival notes from this meeting drew explicit attention to the deteriorating security situation, particularly the increasingly precarious condition of Hindus within the state as a result of Razakar activity (Government of India, Ministry of States, 1947).

In response, Menon proposed a treaty arrangement as a temporary expedient designed to “buy breathing time.” The proposal was not intended as a final settlement but as a stabilizing mechanism that could arrest further deterioration while preserving space for negotiation. Sardar Vallabhbhai Patel, deeply sceptical of Hyderabad's intentions, nonetheless agreed to this approach, as did Prime Minister Jawaharlal Nehru, reflecting a final collective effort by the Indian leadership to prioritize

constitutional resolution over coercion (Government of India, Ministry of States, 1947). A draft agreement was accordingly prepared.

Significantly, the initial draft submitted by Hyderabad was rejected by the Government of India, as it sought to entrench exceptional sovereignty without corresponding obligations. Menon therefore, undertook the unusual step of drafting an alternative agreement on Hyderabad's behalf—an act that underscores the extent to which Indian officials were attempting to rescue negotiations despite diminishing confidence. This revised draft was accepted by Patel, Nehru, and Mountbatten, marking what may be regarded as the last unified constitutional initiative of the Indian leadership toward Hyderabad (Government of India, Ministry of States, 1947). On 26 October 1947, the Hyderabad delegation formally agreed that the draft would be signed by the Nizam and submitted to New Delhi. Yet even at this critical juncture, the delegation was physically surrounded prior to its departure, and the Nizam instructed its members not to proceed—an episode that dramatically illustrated the extent to which internal pressures had overridden diplomatic commitments (Government of India, Ministry of States, 1947).

The internal dynamics behind this reversal became explicit on 28 October 1947, when the Nizam convened all delegation members alongside Kasim Razvi, the leader of the Ittehad-e-Muslimeen. Razvi reportedly assured the Nizam that the Indian Union, already embroiled in crises elsewhere, would be unable to enforce its will should Hyderabad remain defiant (Government of India, Ministry of States, 1947). This assessment was directly shaped by concurrent developments in Kashmir, where Indian troops had been airlifted to Srinagar on 26 October 1947 following the tribal invasion that began on 23 October. The simultaneity of these crises fostered a strategic miscalculation within Hyderabad's ruling circles: that India's attention and military capacity were overstretched, creating a window for defiance.

Acting on this assumption, the Nizam dispatched two emissaries to Karachi on 29 October 1947. By 31 October, Sultan Ahmed informed Menon that Hyderabad was on the verge of agreeing with Pakistan (Government of India, Ministry of States, 1947). The appointment of a new delegation—comprising Nawab Moin Nawaz Jung, Abdur Rahim, and Pingle Venkatarama Reddy—further aggravated Indian suspicions. Sardar Patel's irritation deepened, while Mountbatten expressed acute displeasure upon the delegation's arrival in Delhi on 31 October 1947. Negotiations collapsed definitively on 7 November 1947, culminating in the resignation of the Nawab of Chhatari and his replacement by Mir Laik Ali (also known as Zahid Hussain) (Government of India, Ministry of States, 1947). Around the same time, Kasim Razvi visited Delhi and bluntly informed Menon that “the sword would be the final arbiter,” a statement that crystallized Indian fears that Hyderabad was drifting irreversibly toward confrontation.

Despite this charged atmosphere, a final constitutional arrangement was concluded in the form of the Standstill Agreement, signed by the Nizam on 29 November 1947, along with a collateral letter. On the same day, Sardar Patel laid the agreement before the Constituent Assembly, then functioning as India's provisional Parliament, emphasizing that it represented a temporary expedient rather than a substantive settlement (Government of India, Ministry of States, 1947). The appointment of K. M. Munshi as Agent-General in Hyderabad on 30 November 1947 was intended to monitor compliance and maintain a fragile channel of communication.

The Standstill Agreement, however, proved inherently unstable. On 24 December 1947, a Hyderabad delegation demanded the withdrawal of Indian troops and the transfer of cantonments. While India requested time until April 1948 to implement these changes, the Nizam simultaneously issued two ordinances restricting exports and declaring Indian currency illegal—actions that Menon immediately identified as violations of the agreement (Menon to Hyderabad Government, 25 December 1947, Government of India, Ministry of States, 1947). Hyderabad's evasive response, combined with its offer of a ₹20 crore loan to Pakistan and the appointment of a Public Relations Officer in Karachi, reinforced Indian perceptions of duplicity and strategic drift.

Negotiations slowed following Mahatma Gandhi's assassination on 30 January 1948, even as Razakar violence intensified across Hyderabad's countryside. At a conference of the Premiers of Madras, Bombay, and the Central Provinces on 21 February 1948, presided over by Patel, it was made explicit that while formal accession was not the immediate demand, the Government of India would not compromise on the principle of responsible government within Hyderabad (Government of India, Ministry of States, 1948). This marked a shift from legal formalism to normative legitimacy as the central criterion of settlement.

Subsequent meetings failed to restore confidence. On 2 March 1948, Monckton, Laik Ali, and Nawab Moin Nawaz Jung met Mountbatten, but Menon rejected any supplementary agreements and insisted on concrete restoration of law and order. Patel refused to accept further documentation, signalling a decisive hardening of India's position (Government of India, Ministry of States, 1948). Patel's severe heart attack on 5 March 1948 temporarily slowed decision-making but did not alter strategic direction. Menon's warning letter of 23 March 1948, urging strict adherence to the Standstill Agreement, was met with a lengthy and evasive reply from Laik Ali on 5 April 1948 (Government of India, Ministry of States, 1948).

Even as discussions continued, Razvi openly called for armed conflict on 31 March 1948, underscoring the irrelevance of constitutional assurances. Although a new "Heads of Agreement" was drafted and accepted by India on 26 May 1948, amendments proposed by Laik Ali in early June systematically diluted its substance. Menon reluctantly accepted minor changes, but Patel rejected the revised document

outright on 13 June 1948. Mountbatten, preparing to depart India on 21 June, persuaded Patel to accept the agreement as a final concession. Yet Hyderabad again sought further modifications, which India accepted on 14 June in a last effort to avert conflict (Government of India, Ministry of States, 1948).

The effort proved futile. On 17 June 1948, Monckton informed Mountbatten with a single word—“lost”—signalling the definitive collapse of negotiations. That evening, Nehru publicly declared that India would not move beyond the agreed draft. With Mountbatten's departure and C. Rajagopalachari's assumption of office, the political environment hardened further.

By late August 1948, Hyderabad sought to internationalize the dispute by approaching the United Nations. Despite India's objections, a delegation travelled via Karachi to present its case to the Security Council (Government of India, Ministry of States, 1948). Rajagopalachari's letter of 31 August 1948, urging the Nizam to disband the Razakars, went unanswered. Subsequent correspondence in early September—marked by Hyderabad's assurances and India's detailed documentation of Razakar raids—failed to bridge the divide.

On 9 September 1948, after exhausting diplomatic, constitutional, and political alternatives, Sardar Vallabhbhai Patel authorized Indian military intervention to restore peace and order within Hyderabad. The decision marked not the abandonment of negotiation, but its terminal consequence: the recognition that sovereignty, when persistently decoupled from responsibility, legitimacy, and civilian protection, could no longer be sustained through dialogue alone.

7 THE CONVERGENCE OF WARS AND THE CLOSING OF OPTIONS

Muhammad Ali Jinnah died on 9 September 1948, having repeatedly warned that the threat to the Nizam's state would provoke the uprising of “a hundred million Muslims.” No such uprising occurred. The prophecy collapsed as abruptly as the political order that had sustained it. Whether by design or coincidence, Sardar Vallabhbhai Patel issued the decisive marching order within forty-eight hours of Jinnah's death, signalling that the Government of India no longer regarded restraint as either necessary or prudent.

The decision was communicated to the Southern Command with instructions to prepare for movement in the early hours of 13 September 1948. Even after the date and time were finalized, last-minute efforts were made to postpone the operation until 15 September, reflecting lingering concerns about international reaction and military overstretch. Ultimately, the order stood. Major General J. N. Choudhary was placed in

operational command under the overall direction of Lieutenant General Maharaj Shri Rajendrasinghji, then General Officer Commanding-in-Chief, Southern Command.

The Hyderabad operation unfolded against the backdrop of another unresolved war. As Indian forces prepared to move south, the conflict in Jammu and Kashmir was approaching the completion of its first year—a protracted and inconclusive struggle that had already consumed immense political and military capital. The simultaneity of these two crises was not incidental; it shaped calculations in Delhi, Karachi, and London alike.

The Kashmir conflict began dramatically on 22 October 1947, when Pathan tribal raiders crossed into the state. Two days later, Maharaja Hari Singh appealed to India for military assistance. The Government of India dispatched V. P. Menon to Srinagar for an on-the-spot assessment. Menon advised the Maharaja to evacuate to Jammu and return only after the situation stabilized. On 26 October, Menon reported back to the Cabinet Committee on Defence—attended by Nehru, Patel, and Sheikh Abdullah—where both he and the Prime Minister of Jammu and Kashmir urged immediate military intervention. Lord Mountbatten insisted that accession must precede action. Menon flew to Jammu, found the palace “in utter turmoil,” woke the exhausted Maharaja, secured his signature on the Instrument of Accession, and returned to Delhi the same day.

Indian troops were flown into Srinagar within hours. Yet what followed was not a swift resolution but a grinding war. Acting on Mountbatten's advice, India approached the United Nations on 1 January 1948. Hostilities resumed in earnest in May once the snows melted. Despite tactical gains such as the capture of Tithwal, Indian forces encountered stiff resistance en route to Muzaffarabad. A UN resolution on 13 August 1948 called on Pakistan to vacate the aggression and India to reduce forces pending a plebiscite. Neither side complied. The war ended only on 1 January 1949, with a UN-sponsored ceasefire that froze rather than resolved the conflict.

It was against this unresolved military backdrop that the Hyderabad decision was taken. On the eve of the police action, a high-level conference attended by Nehru, Patel, the three Service Chiefs, and General K. M. Cariappa was held in New Delhi. The stakes were unmistakable: India would now prosecute two wars—one overt, one internal—within a single strategic year.

8 RESISTANCE, OBSTRUCTION, AND STRATEGIC DECEPTION

Internal opposition to the Hyderabad operation came most sharply from Lieutenant General Sir Roy Bucher, Commander-in-Chief of the Indian Army. On 12 September 1948, Bucher strongly advised postponement, arguing that Hyderabad should be dealt with only after Kashmir was settled. He went so far as to threaten resignation if his

advice were ignored. The following day, he attempted to dissuade Patel by warning—without factual basis—that the Hyderabad Air Force might bomb Delhi and Bombay. In reality, Hyderabad possessed no meaningful air capability beyond three transport aircraft operated by Australian gun-runner Sydney Cotton. Patel dismissed the claim outright.

Bucher's scepticism was not limited to Hyderabad. Earlier, on 13 February 1948, he had privately expressed doubts about India's ability to hold the Punch–Mirpur sector in Kashmir, citing troop shortages and logistical constraints. He conveyed the impression that Indian policy aimed merely at extracting refugees rather than securing territory. These views aligned with a broader pattern in which senior British officers continued to shape Indian military thinking in ways that constrained decisive action.

By late 1947, Nehru had already discovered that General Lockhart, Bucher's predecessor, had possessed early intelligence on the tribal invasion of Kashmir but failed to share it with the Government of India. Lockhart's resignation followed, and Bucher assumed command. In January 1948, Cariappa replaced Dudley Russell as commander of the Delhi and East Punjab Command, marking a gradual—though incomplete—Indianization of senior military leadership.

Meanwhile, Hyderabad's military capacity remained modest: six infantry battalions, two light armoured regiments, one horsed cavalry regiment, a field battery, and approximately 1,500 armed irregulars. Leadership quality and fighting efficiency were known to be poor. Yet Hyderabad's significance lay not in its strength but in the strategic distraction it imposed.

Britain's role during this period was decisive and deeply consequential. In September 1948, Pakistan dispatched its Deputy Chief of Army Staff, Major General Cawthorn, to London on a secret mission to negotiate a defence understanding. The mission was warmly received by Prime Minister Clement Attlee. In return, Britain provided Pakistan with “hints” and “tips” regarding Indian intentions. By keeping the Hyderabad issue alive, British policymakers effectively tied down substantial Indian military assets—infantry, armour, and air power—that could otherwise have transformed the balance in Kashmir.

Indeed, over five infantry divisions and the bulk of India's armour and air squadrons were retained away from the Kashmir front. Of twelve armoured regiments, only two were initially deployed there. Air Force squadrons were held at Poona and Kanpur against Hyderabad and East Pakistan, respectively. The war in Kashmir was rendered unwinnable by design.

9 OPERATION POLO AND THE COLLAPSE OF THE NIZAM'S STATE

Against this background, Operation Polo was launched. The plan envisaged a two-pronged advance: the main force moved along the Sholapur–Hyderabad road (186 miles), while a subsidiary column advanced from Bezwada (160 miles). Resistance was encountered on the first two days but rapidly collapsed. On the evening of 17 September 1948, the Hyderabad army surrendered. Major General Choudhuri entered Hyderabad city on 18 September, completing an operation that had lasted 108 hours.

The swiftness of victory owed much to a crucial contingency. On the first day, Indian troops captured Lieutenant T. T. Moore, an ex-British Army commando employed by Hyderabad since August 1947. Moore was driving a jeep loaded with explosives, tasked with demolishing bridges at Naldrug and elsewhere. His papers revealed that Hyderabad expected the Indian advance on 15 September. Had the operation been delayed, bridges would have been destroyed and monsoon rains would have bogged down Indian armour. The decision to advance on the 13th proved decisive.

Indian casualties numbered forty-two, while approximately two thousand Razakars were killed. Within five days, the Indian Army exercised full control over the state. Kasim Razvi was arrested on 19 September and later sentenced to imprisonment. On 23 September, the Nizam formally withdrew Hyderabad's case from the UN Security Council.

The question of Hyderabad's future governance now arose. Menon advised Patel that immediate abolition of the Nizam's dynasty would destabilize Muslim opinion domestically and abroad. Patel agreed, subject to Nehru's concurrence. Nehru also endorsed the approach. Menon travelled to Hyderabad with H. M. Patel and Shavax A. Lal, securing the Nizam's assurance of loyalty. D. S. Bakhle, ICS, assumed office as civil administrator alongside the Military Governor.

Military administration under Major General Choudhuri continued until December 1949, when M. K. Vellodi, ICS, was appointed Chief Minister. A democratic transition followed. In March 1952, general elections were held, and a Congress ministry under B. Ramakrishna Rao assumed office. Hyderabad's integration was complete—not merely through force, but through a calibrated transition from coercion to constitutional normalcy.

10 CONCLUSION

The integration of Hyderabad into the Indian Union emerges from this study not as a singular act of military intervention but as the terminal phase of a prolonged and structurally conditioned process of constitutional exhaustion, administrative consolidation, and sovereign redefinition. By reconstructing the chronology through systematic engagement with archival records of the Ministry of States, Hyderabad

Branch, this article demonstrates that Operation Polo was neither abrupt nor inevitable at the outset, but rather the outcome of a cumulative breakdown of legal, political, and administrative mechanisms that had initially been designed to accommodate princely autonomy within a federal constitutional framework.

The archival evidence reveals that from mid-1947 onward, the Government of India consistently privileged negotiated settlement and constitutional incorporation over coercion. Early correspondence and negotiation files show repeated efforts to secure Hyderabad's accession through treaty arrangements, provisional agreements, and phased constitutional acceptance, even as the geographical and demographic implausibility of independence was acknowledged by Indian negotiators (Ministry of States, Hyderabad Branch, Nizam Property Files, F 57-H/49, NAI). These initiatives reflected a broader postcolonial strategy aimed at preserving legal continuity while dissolving dynastic sovereignty through institutional means rather than force.

However, the archival record also makes clear that this constitutional pathway was progressively undermined from within Hyderabad itself. Files relating to the creation of trusts, management of private property, and foreign financial disclosures indicate an increasing externalization of authority and resources by the Nizam's administration, which complicated the distinction between private rights and sovereign prerogatives (Ministry of States, Hyderabad Branch, Question Arising out of the Creation of Trust by the Nizam, (60)-H/1949, NAI; Declaration of Foreign Holdings Abroad, 7.1(133)-H/1949, NAI). Rather than stabilizing Hyderabad's position, these measures contributed to Indian perceptions that the state was attempting to preserve sovereignty through legal fragmentation and internationalization, rather than through responsible internal governance.

This trajectory became particularly pronounced with evidence of Hyderabad's financial and diplomatic overtures beyond the Indian Union. The archival file concerning the ₹20 crore loan to Pakistan provides critical insight into the security recalibration that followed such actions, marking a decisive shift in Indian policy assessments (Ministry of State, Hyderabad Branch, Hyderabad: 20 Crore Loan to Pakistan, 2(2)-H/1949, NAI). From the perspective of the Government of India, these developments indicated not merely political defiance but a material challenge to territorial and fiscal sovereignty at a moment of fragile post-Partition consolidation.

The post-accession files further illuminate how sovereignty was ultimately reconstituted through institutional normalization rather than dynastic preservation. Records on the fixation of the privy purse under Article 291 of the Constitution demonstrate the transformation of princely authority into a regulated constitutional entitlement, subject to parliamentary and administrative oversight (Ministry of States, Hyderabad Branch, Arrangement of Privy Purse and Private Property of the Nizam; Fixation of Privy Purse under Article 291, 2(1)-H/1950, NAI). Similarly, files documenting the Nizam's acceptance of the draft Constitution underscore the formal

closure of Hyderabad's exceptional status and its incorporation into a uniform constitutional order (Ministry of States, Hyderabad Branch, Provisional Accession of Hyderabad to the Union of India and Acceptance of the Draft Constitution, 1(64)-Hyderabad/1949, NAI).

Methodologically, this study demonstrates the value of AI-assisted analytical tools in historical research, particularly when dealing with heterogeneous archival corpora. Artificial intelligence was employed not as a substitute for historical interpretation but as an organizational and heuristic aid to extract, cluster, and sequence themes across diverse documentary genres—constitutional correspondence, financial records, trust deeds, diplomatic files, and administrative memoranda. This approach enabled the identification of recurring decision points and escalation thresholds that are not readily visible through linear narrative reconstruction alone. By mapping these patterns, the study reveals a rule-like progression in which negotiation repeatedly failed under conditions of internal radicalisation, externalisation of authority, and erosion of responsible governance, culminating in coercive intervention.

Importantly, the AI-assisted framework remains anchored in verifiable primary sources. The archival table and chronology make explicit how each analytical claim is grounded in identifiable files, thereby preserving transparency and historiographical rigor. In this sense, AI functions as a methodological extension of traditional archival practice, enhancing interpretive clarity without displacing contextual judgment or evidentiary standards.

In broader theoretical terms, the Hyderabad case offers a significant contribution to the study of postcolonial state formation. It illustrates how sovereignty in newly independent states was not merely asserted but administratively produced—through negotiations over property, finance, constitutional consent, and international alignment. The failure of Hyderabad's exceptionalism highlights the limitations of legalistic sovereignty, divorced from demographic legitimacy and territorial integration. At the same time, the Indian state's calibrated use of coercion highlights a model of consolidation in which force operated as a corrective mechanism following the exhaustion of constitutional alternatives, rather than as a primary instrument of integration.

In conclusion, the integration of Hyderabad should be understood as a layered historical process in which constitutional negotiation, administrative normalization, and coercive enforcement were sequentially deployed in response to evolving conditions. The combined use of extensive archival evidence and AI-assisted analytical structuring allows this study to move beyond moralized or event-centric interpretations and toward a more systematic understanding of how postcolonial governance systems responded to challenges of fragmented sovereignty. This approach not only refines our understanding of Hyderabad's accession but also offers

a methodological template for analyzing complex episodes of state integration in comparative perspective.

REFERENCES

- Acquisition of Immovable Property by H.E.H. the Nizam of Hyderabad (Bombay), File No. F.10(56)/1949.
- Arrangement of Privy Purse and Private Property of the Nizam; Fixation of Privy Purse under Article 291 of the Constitution of India, File No. 2(1)-H/1950.
- Consent of H.E.H. the Nizam of Hyderabad to Donation of ₹6 Lakhs to the Benaras Hindu University, File No. F.4(20)-H/1949.
- Copland, I. (1988). *The Princes of India in the Endgame of Empire*. Cambridge University Press. [↗](#)
- Dalrymple, W. (2019). *The Anarchy*. Bloomsbury. [↗](#)
- Declaration of Foreign Holdings Abroad, File No. 7.1(133)-H/1949.
- Grant Made by the Nizam to the Jama Masjid of Delhi, File No. 2(1)-H/1949.
- Guha, R. (2007). *India After Gandhi*. HarperCollins. [↗](#)
- Hyderabad: ₹20 Crore Loan to Pakistan, File No. 2(2)-H/1949.
- Menon, V.P. (1956). *The Story of the Integration of the Indian States*. *Orient Longman*. [↗](#)
- Nizam Property Files, File No. F 57-H/49.
- Prince of the Bera Trust, File No. F.1(58)-H/1949.
- Programme of Egyptian Journalist Delegation Visiting India, File No. 12(12)-H/1949.
- Provisional Accession of Hyderabad to the Union of India and Acceptance by the Nizam of the Draft Constitution for the State, File No. 1(64)-Hyderabad/1949.
- Provisional Accession of Hyderabad to the Union of India and Draft Constitution for States Adopted by the Constituent Assembly of India, File No. (64)-H/1949.
- Question Arising out of the Creation of Trust by the Nizam, File No. (60)-H/1949.
- Richards, J.F. (1993). *The Mughal Empire*. Cambridge University Press. [↗](#)
- Sarkar, J. (1983). *History of Hyderabad*. Orient Longman. [↗](#)

Artificial Intelligence as an Enabler of Smart Medical Tourism: A Comparative Study of India and Uzbekistan

Jiyanov Uktam Panjiyevich

*Department of Islamic Economics, Finance, and Pilgrimage Tourism,
International Islamic Academy of Uzbekistan, Tashkent, Uzbekistan*

Nidhi Singh

Amity University, Noida, India

Esonboev Bakhodir Bakir Ugli

*Department of Islamic Economics, Finance, and Pilgrimage Tourism,
International Islamic Academy of Uzbekistan, Tashkent, Uzbekistan*

DOI: [10.1201/9781042018116-12](https://doi.org/10.1201/9781042018116-12)

ABSTRACT: The rapid integration of artificial intelligence (AI) into healthcare systems is reshaping the global landscape of medical tourism by enabling more efficient, transparent, and patient-centred cross-border care. This paper examines AI as a critical enabler of “smart medical tourism” through a comparative study of India and Uzbekistan—two countries occupying distinct yet complementary positions within emerging Eurasian healthcare corridors. India's large-scale health infrastructure, long-standing medical tourism experience, and advanced adoption of digital health platforms position it as a regional hub for AI-supported diagnostics, telemedicine, and hospital management systems. In contrast, Uzbekistan is

progressively leveraging AI within its expanding health infrastructure and natural therapeutic environments, including wellness- and rehabilitation-based tourism, as part of a broader strategy for healthcare modernization and international integration. The study analyses how AI applications—ranging from predictive analytics and clinical decision support to multilingual patient navigation, logistics optimization, and post-treatment follow-up—are being adopted at different stages of maturity in both countries. Beyond operational efficiency, the paper explores how AI enhances trust, accessibility, and continuity of care for international patients while reducing informational asymmetries and transaction costs. It also critically addresses challenges related to data governance, ethical accountability, interoperability, and regulatory divergence. By positioning AI as a decision-support and coordination architecture rather than a replacement for clinical judgment, the article argues that AI-enabled smart medical tourism can foster equitable, resilient, and patient-centric healthcare mobility between India and Uzbekistan, contributing to broader debates on digital health diplomacy and sustainable cross-border cooperation.

Keywords: Artificial Intelligence in Healthcare, Smart Medical Tourism, Health Diplomacy, Healthcare Mobility, and India–Uzbekistan

1 METHODOLOGY

This study adopts a qualitative, comparative, and interdisciplinary methodology to examine the role of artificial intelligence as an enabler of smart medical tourism between India and Uzbekistan. The analysis is based on a systematic review of secondary sources, including government reports, policy documents, international organization publications (such as WHO and ADB), tourism and health statistics, and peer-reviewed literature on AI, digital health, and medical tourism. A comparative framework is employed to assess differences and complementarities across key dimensions—healthcare infrastructure, AI and digital health adoption, medical tourism volume, regulatory readiness, and wellness integration. Conceptual tools

such as maturity indices, trend analysis, and illustrative comparative figures are used to synthesize data and highlight patterns rather than to produce econometric measurement. The study is further informed by normative and critical analysis, emphasizing ethical governance, data protection, and human-centred AI, thereby positioning AI as a decision-support and coordination architecture within cross-border healthcare mobility rather than as a substitute for clinical judgment.

2 INTRODUCTION

Artificial intelligence is increasingly recognized as a transformative force in the global medical tourism sector, enhancing efficiency, patient experience, and care coordination across borders ([IndiaAI.gov.in, 2025](#)). In the context of a comparative study of India and Uzbekistan, AI functions as a critical enabler of “smart medical tourism” by integrating advanced data analytics, decision support systems, and digital patient services into the healthcare value chain ([IndiaAI.gov.in, 2025](#)). Smart medical tourism refers to the use of digital and AI-driven technologies to streamline medical travel, ranging from personalized treatment mapping and risk prediction to multilingual patient navigation and telemedicine support.

India has emerged as a major hub for medical tourism due to its broad clinical capacities, internationally accredited hospitals, and expanding adoption of digital health technologies. India attracts hundreds of thousands of foreign patients annually—including patients from Uzbekistan—because of its cost-effective, high-quality care and technology-enabled service platforms ([Bureau of Immigration, 2025](#); [Medijourney, 2025](#)). AI is already being integrated into Indian healthcare through predictive analytics, telehealth services, and platforms that facilitate remote assessment and cross-border patient engagement, thereby enhancing the attractiveness of India as a smart medical tourism destination ([IndiaAI.gov.in, 2025](#)). Partnerships between Indian medical associations and global digital health platforms further illustrate how AI is being systematically used to improve international patient referrals, remote triage, and treatment outcomes ([ETHealthWorld, 2025](#)).

Uzbekistan's healthcare system, while still developing, has shown rapid growth in medical tourism capacity, as reflected in increasing numbers of foreign patients seeking treatment within the country and broader healthcare improvements ([National Statistics Committee of Uzbekistan, 2025](#)). Although Uzbekistan's adoption of AI in healthcare is less documented than India's, the country has taken steps to enhance healthcare services and expand its appeal as a medical tourism destination, including improvement of medical institutions to meet international standards and growth in foreign medical visitors ([National Statistics Committee of Uzbekistan, 2025](#)). Bilateral cooperation in the healthcare sector—including hospital investments, exchange programs, and joint training—creates a foundation for shared innovation where AI tools could support quality enhancement, workforce training, and operational coordination ([UzA, 2024](#); UzDaily, 2025).

Together, India and Uzbekistan occupy *distinct yet complementary positions* in the emerging Eurasian healthcare corridor: India as an established global medical tourism hub with sophisticated AI adoption in digital health infrastructure, and Uzbekistan as a growing regional destination investing in healthcare improvements and intergovernmental collaboration. This complementary dynamic creates opportunities to leverage AI not only for improving clinical outcomes and operational efficiencies but also for developing an interconnected, smart medical tourism ecosystem that facilitates seamless patient journeys across borders.

3 AI, DIGITAL HEALTH, AND THE TRANSFORMATION OF MEDICAL TOURISM

India's contemporary leadership in global medical tourism represents a historical continuum extending from its ancient traditions of healing to its present embrace of artificial intelligence (AI) and digital health technologies. While India's civilizational medical ethos emphasized ethical responsibility, compassion, and professional duty, its modern healthcare system has increasingly integrated AI-enabled diagnostics, telemedicine, hospital information systems, and digital patient platforms to enhance

cross-border care delivery ([World Health Organization \[WHO\], 2021](#)). This convergence of tradition and technology has positioned India as a leading destination for “smart medical tourism,” where digital infrastructures support efficiency, transparency, and patient-centred outcomes.

AI has become a critical enabler across the medical tourism value chain. In Indian tertiary-care hospitals, AI-assisted imaging, clinical decision-support systems, and robotic surgery enhance precision and treatment outcomes in high-demand specialties such as cardiology, oncology, orthopaedics, and organ transplantation ([NITI Aayog, 2022](#)). Telemedicine platforms supported by AI-driven triage systems allow international patients to consult Indian specialists prior to travel, reducing uncertainty and improving clinical preparedness ([Ministry of Health and Family Welfare \[MoHFW\], 2023](#)). Post-treatment continuity of care is further supported through interoperable electronic health records and remote monitoring tools, transforming medical tourism from episodic intervention into a digitally sustained care continuum.

Government initiatives such as *Heal in India* and the promotion of Medical Value Travel (MVT) increasingly emphasise digital readiness and platform-based coordination. The Indian medical tourism market, valued at USD 2.89 billion in 2020, is projected to reach USD 13.42 billion by 2026, driven in part by AI-enabled efficiency, reduced waiting times, and improved international patient navigation ([Ministry of Tourism, 2022](#)). AI-powered multilingual interfaces and digital concierge services further reduce informational asymmetries, a critical barrier for international patients navigating unfamiliar healthcare systems ([UN World Tourism Organization \[UNWTO\], 2021](#)).

Uzbekistan's growing participation in India's medical tourism ecosystem illustrates the emergence of AI-enabled Eurasian healthcare corridors. The number of Uzbek patients travelling to India for medical treatment increased from 1,712 in 2020 to 8,921 in 2024, reflecting rising trust in India's healthcare infrastructure and digital health capabilities ([Press Information Bureau \[PIB\], 2024](#)). According to the Embassy of India in Tashkent, approximately 8,000 Uzbek nationals now seek treatment in India annually, particularly for advanced procedures in cardiology, oncology, and orthopaedics.

While India represents a mature medical tourism hub with advanced AI adoption, Uzbekistan occupies a complementary position as a rapidly integrating partner seeking healthcare modernisation through international collaboration. Uzbekistan has prioritised healthcare reforms, digital governance, and foreign partnerships to enhance service quality and accessibility ([Asian Development Bank \[ADB\], 2022](#)). AI-enabled teleconsultations and digital referral platforms reduce travel risks for Uzbek patients and facilitate coordinated care pathways with Indian hospitals. This asymmetrical yet complementary relationship enables India to export medical expertise and digital health services, while Uzbekistan benefits from capacity-building, training, and access to specialised care.

The ethical integration of AI into India's medical tourism sector is deeply informed by its civilizational understanding of medicine. Classical texts such as the *R̥gveda*, *Charaka Saṃhitā*, and *Suśruta Saṃhitā* conceptualised health as a holistic balance of body, mind, and spirit, emphasizing compassion, restraint, and accountability in medical practice ([Sharma, 2015](#)). In this tradition, technology serves as a means to healing rather than an autonomous authority. Contemporary AI governance in Indian healthcare increasingly reflects this ethos by positioning AI as a decision-support architecture that augments clinical judgment rather than replacing human agency ([NITI Aayog, 2022](#)).

This alignment between ancient ethical principles and modern digital governance distinguishes India's approach to AI-enabled medical tourism. By embedding transparency, explainability, and patient welfare into AI deployment, India offers a model of humane digital health innovation that balances efficiency with empathy ([WHO, 2021](#)). The integration of artificial intelligence and digital health has transformed India's medical tourism sector into a globally competitive, technologically advanced, and ethically grounded ecosystem. The growing participation of Uzbekistan underscores how AI-enabled healthcare mobility is reshaping transregional cooperation, health diplomacy, and people-to-people engagement. Together, India and Uzbekistan occupy distinct yet complementary positions within emerging Eurasian healthcare corridors—India as a digital-health and clinical hub, and Uzbekistan as a rapidly integrating partner. This evolving relationship demonstrates how smart medical tourism, grounded in both

technological innovation and civilizational ethics, can contribute to sustainable, equitable global healthcare futures.

4 INDIA'S AI-ENABLED HEALTH INFRASTRUCTURE AS A REGIONAL MEDICAL TOURISM HUB

India's emergence as a premier regional hub for AI-supported medical tourism rests on three interlocking assets: a large-scale health infrastructure, decades of experience in international healthcare services, and the advanced adoption of digital health platforms powered by artificial intelligence.

First, India's health ecosystem incorporates a vast network of tertiary-care hospitals, specialised clinics, and accredited institutions capable of delivering complex procedures at scale. With thousands of hospitals certified by national and international bodies, India offers comprehensive services in cardiology, oncology, orthopaedics, organ transplantation, and critical care—many of which are increasingly augmented by AI-driven diagnostics and clinical decision support systems ([NITI Aayog, 2022](#)). AI-assisted imaging tools, machine-learning models for risk stratification, and predictive analytics modules help clinicians interpret large datasets, improve diagnostic accuracy, and reduce time-to-treatment, directly addressing the needs of international patients seeking high-quality care ([World Health Organization \[WHO\], 2021](#)).

Second, India's long-standing experience in medical tourism has created institutional and procedural familiarity with cross-border healthcare mobility, including streamlined administrative processes, dedicated medical visa categories, and well-developed patient concierge services. This foundation allows digital innovations—such as AI-enabled multilingual patient interfaces, remote triage platforms, and real-time care coordination systems—to scale rapidly in response to international demand ([Ministry of Tourism, 2022](#)). AI-driven telemedicine platforms further extend India's reach, allowing prospective patients in remote and international settings to engage with specialists long before travel, thereby increasing global confidence and reducing barriers to access ([MoHFW, 2023](#)).

Finally, India's digital health revolution—embodied by initiatives like the Ayushman Bharat Digital Mission and interoperable electronic health records—enables seamless integration of AI across hospital management systems, patient portals, and telehealth services ([Ministry of Health and Family Welfare, 2023](#)). These platforms support resource optimisation, appointment scheduling, inventory control, and predictive workforce planning while ensuring continuity of care across regional and international patient cohorts. By integrating AI into everyday clinical and administrative workflows, India not only enhances internal efficiency but also elevates the quality of care offered to foreign patients, reinforcing its role as a regional center for smart medical tourism.

These strengths position India to collaborate with neighbouring and partner countries such as Uzbekistan, enabling shared digital health pathways, interoperable telemedicine linkages, and AI-supported referrals that can improve diagnostic confidence and continuity of care across borders. In contrast to India's established medical tourism ecosystem, Uzbekistan is progressively integrating AI and digital health innovation within an expanding healthcare infrastructure, partly as a strategic response to both domestic healthcare needs and international integration objectives.

Uzbekistan's national healthcare strategy has prioritised modernisation, digital transformation, and international cooperation to raise service quality and expand access ([Asian Development Bank \[ADB\], 2022](#)). While large-scale AI adoption is still emerging, Uzbekistan has begun deploying AI-assisted teleconsultations, remote diagnostics, and predictive analytics in selected tertiary facilities, particularly in urban centres, as part of efforts to integrate international best practices and enhance clinical capacity. This approach aligns Uzbekistan's healthcare growth with global digital standards, enabling clinicians to access decision support tools that improve diagnostic accuracy and guide treatment planning—especially for complex conditions where specialist expertise may be limited locally.

Moreover, Uzbekistan is leveraging its natural therapeutic environments—including mineral springs, rehabilitative centres, and wellness resorts—in combination with digital and AI tools to develop wellness- and rehabilitation-based tourism. By integrating wearable health monitors, AI-mediated patient progress tracking, and personalised care algorithms, these facilities are expanding beyond traditional spa tourism into data-driven,

health-oriented ecosystems. Such convergence of natural therapeutic assets with digital health platforms reflects a broader strategy to diversify Uzbekistan's appeal to international visitors while strengthening domestic healthcare outcomes.

Uzbekistan's growing medical travel links with India illustrate this strategic trajectory. The number of Uzbek patients seeking treatment in India has increased significantly, with approximately 8,921 arriving in 2024, up from 1,712 in 2020 ([Press Information Bureau \[PIB\], 2024](#)). The increasing use of AI-enabled telemedicine consultations and pre-travel digital assessments helps Uzbek patients navigate the clinical landscape of Indian hospitals, reducing uncertainty and enhancing cross-border care coordination. These interactions, in turn, expose Uzbek health professionals and policymakers to Indian digital health practices, fostering bilateral knowledge exchange and capacity building.

This progressive adoption of AI—matched with investments in wellness tourism and health infrastructure—positions Uzbekistan as a complementary partner in a broader Eurasian healthcare corridor. While India contributes deep experience, advanced clinical infrastructure, and mature digital ecosystems, Uzbekistan brings strategic modernization goals, natural therapeutic assets, and a growing commitment to AI integration that collectively advance shared objectives in regional health cooperation. Together, India and Uzbekistan exemplify how AI and digital health technologies can drive bilateral progress in medical tourism and healthcare modernization. India's expansive health infrastructure and advanced digital platforms offer scalable models of AI-enabled clinical and administrative excellence, while Uzbekistan's strategic adoption of AI and emphasis on wellness tourism expand the scope of cross-border healthcare engagement. Through interoperable digital linkages, telemedicine collaborations, and shared innovation platforms, both countries can progress in tandem toward a more integrated, efficient, and patient-centred regional healthcare landscape.

5 DIFFERENTIAL ADOPTION AND IMPACT OF AI APPLICATIONS IN INDIA AND UZBEKISTAN

Artificial intelligence (AI) is increasingly embedded across the medical tourism value chain, from predictive analytics and clinical decision support to patient navigation, logistics optimization, and post-treatment follow-up. However, the stage of adoption and institutional maturity of these applications varies significantly between India and Uzbekistan, reflecting each country's healthcare priorities, digital infrastructure, and integration strategies.

In India, AI applications have reached higher maturity across multiple clinical and administrative domains, driven by robust digital health policies, substantial investment in health IT infrastructure, and a competitive private healthcare sector. At the clinical front end, predictive analytics and machine learning models are widely used to support early disease detection and risk stratification. For example, AI-assisted imaging algorithms enhance accuracy in radiology and pathology by identifying subtle diagnostic patterns that may escape human interpretation, reducing diagnostic delays and improving prognostic outcomes ([World Health Organization \[WHO\], 2021](#)). Clinical decision-support systems embedded within electronic health records (EHRs) further assist physicians by suggesting evidence-based treatment pathways and alerting providers to potential medication interactions or adverse events ([NITI Aayog, 2022](#)).

Beyond diagnostics, multilingual patient navigation platforms powered by natural language processing (NLP) and AI chatbots are increasingly deployed to help international patients interpret medical information, schedule appointments, and clarify procedural steps. These systems reduce cultural and linguistic barriers that often impede care access, enhancing patient confidence and autonomy in destination healthcare settings ([Ministry of Health and Family Welfare \[MoHFW\], 2023](#)).

AI-driven logistics optimization tools play a crucial role in coordinating cross-border patient flows, hospital resources, and care transitions. Algorithms designed for predictive demand forecasting enable hospitals to anticipate patient volumes and allocate resources efficiently, while AI-enabled scheduling systems ensure synchronized coordination of appointments, tests, and procedures across departments. These functions reduce waiting times and minimize costs associated with inefficiencies—benefits that are especially valuable to international patients managing complex treatment itineraries.

Finally, post-treatment follow-up is increasingly supported by wearable sensors, remote monitoring platforms, and AI-based alert systems that detect deviations from expected recovery patterns. These innovations facilitate continuity of care after discharge, ensuring that patients remain connected with care teams and receive timely interventions if complications arise ([MoHFW, 2023](#)). Collectively, these AI applications enhance trust, accessibility, and seamless care continuity, while reducing informational asymmetries and transaction costs for foreign patients navigating a foreign healthcare ecosystem.

6 UZBEKISTAN: EMERGING AI ADOPTION AND STRATEGIC MODERNIZATION

In Uzbekistan, AI adoption in healthcare is at an early yet rapidly expanding stage, reflecting the government's broader strategy of healthcare modernization and international integration ([Asian Development Bank \[ADB\], 2022](#)). While core medical facilities are strengthening their clinical infrastructure, pilot implementations of predictive analytics and decision support tools are beginning to emerge in select tertiary centres. These early AI initiatives often focus on risk forecasting and hospital management applications, allowing administrators to streamline patient triage and optimize resource utilization in busy urban hospitals. Although not yet as widespread as in India, these nascent systems reflect an intentional alignment with international digital health norms.

Uzbekistan has also begun experimenting with AI-enabled telemedicine platforms that provide remote consultations and preliminary assessments. These platforms are particularly valuable for patients in rural regions or those with limited access to specialised care, helping bridge geographic disparities and increasing healthcare accessibility ([ADB, 2022](#)). The integration of AI translation modules and simplified navigation tools further helps patients orient themselves within evolving digital health interfaces, building initial layers of trust and reducing linguistic friction.

In the realm of logistics, Uzbekistan's AI adoption is currently modest but strategically targeted. Health administrators are exploring AI-assisted

scheduling and patient flow management systems to reduce bottlenecks in outpatient clinics and diagnostic facilities. Similarly, early models of post-treatment tele-follow-up systems are being tested, often supported by wearable devices and mobile apps that enable remote monitoring and patient engagement post-discharge.

Though Uzbekistan's AI infrastructure remains in formative stages relative to India, these initial applications are designed to incrementally enhance patient accessibility, reduce administrative friction, and strengthen continuity of care. As these systems mature, they are expected to reduce transaction costs—both financial and cognitive—for patients navigating the healthcare system, especially when cross-border engagement requires coordination with external providers such as Indian hospitals.

Across both contexts, AI plays a pivotal role in addressing three key systemic challenges that traditionally hinder international patient mobility: trust deficits, informational asymmetries, and transaction costs. Trust is reinforced through transparent AI-enabled decision-support and communication tools that provide consistent, evidence-based insights into diagnoses, treatment options, and expected outcomes. Accessibility improves as AI platforms reduce language barriers, support remote engagement, and automate complex scheduling and coordination tasks that previously required significant human interaction and administrative effort.

AI also reduces informational asymmetries—knowledge gaps between patients and providers that can lead to misinterpretation, fear, or suboptimal decisions—by providing patients with personalised, context-specific information in accessible formats. Predictive analytics supply anticipatory insights that empower both clinicians and patients, while logistics optimization minimizes transactional friction by matching patient needs with available resources efficiently.

For international patients, these advantages translate into lower effective costs, fewer delays, and more predictable care pathways. In India, AI's maturity across clinical and administrative ecosystems yields measurable operational gains that benefit both domestic and international cohorts. In Uzbekistan, emerging AI adoption lays the groundwork for improved user experiences and cross-border healthcare linkages. Together, these developments illustrate how AI can transform national healthcare systems

and create interoperable, patient-focused regional health corridors that support sustainable medical tourism and shared healthcare modernization.

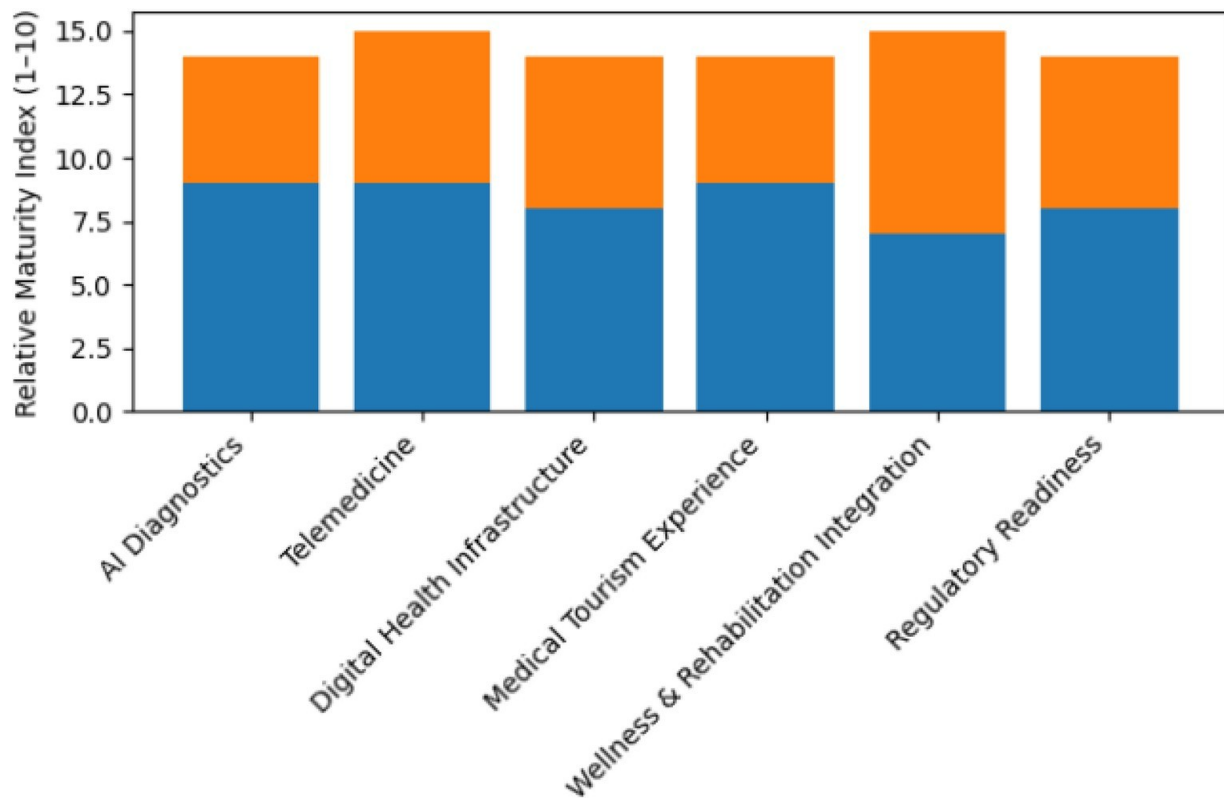


Figure 1. Comparative maturity of AI-Enabled Smart Medical Tourism: India and Uzbekistan. (Note: Blue colour represents India's performance). (Source: Data acquired from various government data sources of India and Uzbekistan).

The comparative graph illustrates a relative maturity index (ranging from 1 to 10) across six strategically significant dimensions of AI-enabled smart medical tourism: AI-supported diagnostics, telemedicine adoption, digital health infrastructure, medical tourism experience, wellness and rehabilitation integration, and regulatory readiness. The visualization highlights clear asymmetries as well as complementarities between India and Uzbekistan. India demonstrates a higher level of maturity in AI-supported diagnostics, telemedicine platforms, digital hospital management systems, and accumulated medical tourism experience, reflecting its large-scale healthcare infrastructure, advanced digital health ecosystem, and long-

standing engagement with international patients. These strengths position India as a regional hub for technologically intensive and clinically complex medical travel. In contrast, Uzbekistan shows relative strength in wellness- and rehabilitation-based healthcare tourism, supported by its natural therapeutic environments and a policy-driven commitment to healthcare modernization. The country's emerging regulatory and digital health initiatives indicate a growing institutional readiness to integrate AI into its evolving healthcare and tourism sectors. Taken together, the figure visually reinforces the central argument of this study: rather than competing directly, India and Uzbekistan occupy distinct yet complementary positions within an emerging Eurasian smart medical tourism corridor, where technological leadership, clinical expertise, and wellness-oriented care can be strategically aligned through AI-enabled coordination.

Governing AI in Cross-Border Medical Tourism: Risks, Ethics, and Strategic Coordination

While artificial intelligence offers substantial promise in enabling smart medical tourism between India and Uzbekistan, its deployment across borders raises complex challenges related to data governance, ethical accountability, interoperability, and regulatory divergence. These challenges are not merely technical but deeply political and normative, shaping how trust, responsibility, and authority are distributed within transnational healthcare systems.

Data governance constitutes one of the most critical fault lines in AI-enabled medical tourism. Cross-border healthcare mobility relies heavily on the movement of sensitive personal health data, including diagnostic images, genomic information, treatment histories, and biometric records. In India, the expansion of digital health platforms under the Ayushman Bharat Digital Mission has improved data standardisation and interoperability but continues to grapple with questions of consent, data ownership, and secondary data use ([Ministry of Health and Family Welfare \[MoHFW\], 2023](#)). Uzbekistan, meanwhile, is in the process of strengthening its digital governance frameworks, with evolving regulations governing data localisation and cybersecurity ([Asian Development Bank \[ADB\], 2022](#)).

The absence of harmonised data protection standards between the two countries introduces risks of data fragmentation, jurisdictional ambiguity, and uneven accountability. Without interoperable yet sovereign-respecting

data governance mechanisms, AI systems trained on cross-border datasets may reinforce asymmetries in control and benefit distribution, particularly disadvantaging patients from countries with less mature regulatory ecosystems ([World Health Organization \[WHO\], 2021](#)). This underscores the need for bilateral or regional data governance protocols that balance innovation with patient rights and national sovereignty.

Ethical accountability in AI-enabled healthcare extends beyond concerns of privacy to questions of bias, explainability, and responsibility for clinical outcomes. AI systems used for diagnostics, risk prediction, or triage inevitably reflect the data on which they are trained. Given that most advanced clinical AI models deployed in India are trained on datasets drawn predominantly from Indian or Western populations, their applicability to Uzbek patients may be limited, raising concerns of algorithmic bias and misclassification ([NITI Aayog, 2022](#)).

From an ethical standpoint, delegating excessive authority to AI systems risks obscuring responsibility in clinical decision-making. If an AI-supported diagnostic recommendation leads to adverse outcomes, determining liability—whether it lies with the physician, the hospital, the software provider, or the data source—becomes legally and morally complex. These ambiguities are magnified in cross-border medical tourism, where legal jurisdictions overlap and patient recourse mechanisms are often unclear ([WHO, 2021](#)). A critical ethical position therefore requires AI to remain transparent, auditable, and subordinate to human clinical judgment, rather than functioning as an autonomous decision-maker.

Interoperability represents another structural challenge to AI-enabled smart medical tourism. India's digital health ecosystem is advancing toward interoperable electronic health records and standardized health identifiers, yet integration across private providers remains uneven ([MoHFW, 2023](#)). Uzbekistan's digital health systems, while expanding, are still characterized by fragmented platforms and limited cross-institutional data exchange ([ADB, 2022](#)). This asymmetry complicates the creation of seamless AI-supported care pathways for international patients.

Without interoperable technical standards, AI systems cannot reliably aggregate or interpret patient data across borders, reducing their effectiveness in longitudinal care coordination and post-treatment follow-up. More critically, interoperability gaps can reinforce institutional silos and

increase transaction costs for patients, undermining the very efficiencies that AI is meant to deliver. Addressing this requires not only technological alignment but institutional

Regulatory divergence between India and Uzbekistan further constrains the deployment of AI in medical tourism. India's AI governance discourse increasingly emphasises “responsible AI,” ethical safeguards, and sector-specific regulation, while Uzbekistan's regulatory framework remains under development as part of broader digital state-building efforts ([NITI Aayog, 2022](#); [ADB, 2022](#)). This divergence creates uncertainty for healthcare providers, technology developers, and patients operating across jurisdictions.

A critical concern here is the risk of techno-solutionism—the assumption that AI alone can resolve systemic healthcare challenges without addressing institutional capacity, workforce training, or ethical oversight. Overreliance on AI-driven efficiency metrics may inadvertently marginalise human care, reduce professional autonomy, and weaken trust if patients perceive technology as replacing rather than supporting physicians ([WHO, 2021](#)). Such outcomes would be particularly detrimental in medical tourism, where trust and reassurance are central to patient decision-making.

In response to these challenges, this paper argues for positioning AI not as a replacement for clinical judgment but as a decision-support and coordination architecture embedded within human-centred governance frameworks. In this model, AI augments physicians’ diagnostic and administrative capacities, enhances coordination across institutions, and supports continuity of care—while leaving final authority with trained medical professionals.

For India and Uzbekistan, this approach enables asymmetric yet equitable cooperation. India contributes advanced AI systems, clinical expertise, and digital platforms, while Uzbekistan engages through contextual adaptation, capacity-building, and the integration of AI into wellness and rehabilitation ecosystems. This configuration allows both countries to progress in tandem without exacerbating dependency or undermining local agency.

By foregrounding ethical governance, interoperability, and shared responsibility, AI-enabled smart medical tourism can become a vehicle for digital health diplomacy rather than a source of regulatory conflict. Bilateral

cooperation on AI standards, data governance, and professional training can strengthen trust, reduce systemic risk, and promote sustainable healthcare mobility. In this sense, AI is not merely a technological tool but a diplomatic infrastructure shaping how states collaborate, share expertise, and protect patient welfare across borders. Critically governed, AI-enabled medical tourism between India and Uzbekistan has the potential to foster equitable, resilient, and patient-centric healthcare mobility, contributing to broader debates on sustainable globalization, ethical digital transformation, and South–South cooperation in health systems.

7 AI, MEDICAL TOURISM, AND THE RECONFIGURATION OF EURASIAN HEALTHCARE MOBILITY

The comparative analysis of India and Uzbekistan demonstrates that artificial intelligence is no longer a peripheral technological add-on in medical tourism but a foundational infrastructure reshaping how cross-border healthcare is imagined, governed, and experienced. As this study has shown, AI-enabled smart medical tourism is transforming medical travel from a fragmented, episodic encounter into a coordinated, data-driven, and patient-centric continuum of care. India and Uzbekistan occupy distinct yet complementary positions within this emerging Eurasian healthcare corridor, and their evolving interaction illustrates both the possibilities and the constraints of AI-driven healthcare mobility.

India's role as a regional hub for smart medical tourism is anchored in the convergence of scale, experience, and digital maturity. Its vast hospital network, long-standing engagement with international patients, and institutional familiarity with medical value travel provide fertile ground for the integration of AI-supported diagnostics, telemedicine, hospital management systems, and digital patient navigation platforms ([Ministry of Tourism, 2022](#); [NITI Aayog, 2022](#)). AI in the Indian context functions as an efficiency multiplier—enhancing diagnostic accuracy, optimizing resource allocation, reducing waiting times, and enabling continuity of care beyond physical borders through teleconsultation and remote monitoring ([World Health Organization \[WHO\], 2021](#)). These capabilities not only strengthen

India's global competitiveness but also position it as a provider of digital health public goods for partner countries seeking modernization without replicating high-cost infrastructure.

Uzbekistan, by contrast, represents a strategically significant but differently situated node in this ecosystem. Its healthcare modernization agenda reflects a deliberate attempt to integrate advanced technologies, international standards, and wellness-based tourism into a coherent national strategy ([Asian Development Bank \[ADB\], 2022](#)). While AI adoption in Uzbekistan remains at an earlier stage of maturity, the country's investments in digital governance, telemedicine, and rehabilitative and wellness infrastructure indicate a future-oriented approach. Importantly, Uzbekistan's natural therapeutic environments—mineral springs, rehabilitation centres, and wellness resorts—create a distinctive platform where AI-enabled personalization, progress tracking, and post-treatment monitoring can complement conventional clinical services. In this sense, Uzbekistan is not merely a recipient of Indian medical expertise but an active participant in diversifying the geography and modalities of smart medical tourism.

The increasing flow of Uzbek patients to India—rising from 1,712 in 2020 to 8,921 in 2024—illustrates how AI-mediated pathways are already shaping patient mobility, trust formation, and healthcare decision-making ([Press Information Bureau \[PIB\], 2024](#)). Teleconsultations, AI-supported diagnostics, and digital referral platforms reduce informational asymmetries and transaction costs that traditionally deter cross-border care, particularly for patients from emerging economies. These technologies transform trust from a purely reputational phenomenon into a data-supported and process-driven relationship, where transparency, predictability, and continuity of care become central to patient confidence ([WHO, 2021](#)).

However, this study also underscores that the promise of AI-enabled medical tourism is inseparable from its governance challenges. Data governance remains a critical vulnerability, particularly in cross-border contexts where health data traverses multiple legal and regulatory regimes. India's expanding digital health architecture has made significant progress in standardization and interoperability, yet concerns regarding consent, secondary data use, and algorithmic accountability persist ([Ministry of Health and Family Welfare \[MoHFW\], 2023](#)). Uzbekistan's evolving regulatory frameworks, while improving, highlight the asymmetries that

can emerge when AI systems are deployed across jurisdictions with uneven institutional capacity ([ADB, 2022](#)). Without harmonized or mutually recognized governance mechanisms, AI risks reproducing digital dependency rather than enabling equitable cooperation.

Ethical accountability presents an equally pressing concern. AI systems trained predominantly on Indian or Western datasets may not adequately reflect the epidemiological, genetic, or socio-cultural characteristics of Uzbek patients, raising the risk of bias and misclassification ([NITI Aayog, 2022](#)). In medical tourism, where patients often operate outside their home legal systems, unclear lines of responsibility for AI-assisted decisions can undermine both trust and justice. This analysis therefore reinforces a critical normative position: AI must remain a decision-support and coordination architecture, not a substitute for clinical judgment. The physician—ethically accountable, culturally situated, and professionally trained—must remain central to care delivery, with AI serving as an augmentative tool rather than an autonomous authority ([WHO, 2021](#)).

Positioning AI in this manner carries important implications for digital health diplomacy. When governed responsibly, AI-enabled medical tourism can function as a diplomatic infrastructure that deepens bilateral cooperation, facilitates knowledge exchange, and aligns healthcare modernization with people-to-people engagement. For India and Uzbekistan, collaboration on AI standards, interoperable digital health platforms, professional training, and ethical governance frameworks offers a pathway toward resilient and sustainable healthcare mobility. Such cooperation moves beyond transactional medical travel toward a shared regional vision of health security, innovation, and inclusivity.

In broader theoretical terms, this study contributes to debates on globalization and technology by demonstrating that AI does not flatten healthcare geographies uniformly. Instead, it reconfigures them through asymmetrical but potentially complementary relationships. India's technological and clinical leadership, when paired with Uzbekistan's modernization goals and wellness assets, illustrates how South–South cooperation can leverage AI without replicating extractive or hierarchical models of digital globalization. The result is not a singular medical tourism destination but an interconnected corridor in which care, data, expertise, and trust circulate dynamically.

In conclusion, AI-enabled smart medical tourism between India and Uzbekistan represents both an opportunity and a responsibility. If pursued uncritically, it risks reinforcing regulatory gaps, ethical ambiguities, and digital inequalities. If governed with foresight, transparency, and human-centred values, it can foster equitable, patient-centric, and resilient healthcare mobility. By embedding AI within ethical traditions of care, robust governance frameworks, and cooperative diplomacy, India and Uzbekistan can jointly shape a model of medical tourism that is not only smart but just—one that aligns technological innovation with the enduring moral imperative of healing across borders.

REFERENCES

- Asian Development Bank. (2022). *Uzbekistan: Health sector assessment*. ADB.[↗](#)
- Ministry of Health and Family Welfare. (2023). *Telemedicine and Digital Health Initiatives in India*. Government of India.[↗](#)
- NITI Aayog. (2022). *National Strategy on Artificial Intelligence: Responsible AI for all*. Government of India.
- World Health Organization. (2021). *Ethics and Governance of Artificial Intelligence for Health*. WHO.[↗](#)
- Asian Development Bank. (2022). *Uzbekistan: Health Sector Assessment*. ADB.
- Bureau of Immigration. (2025). Foreign tourist arrivals data for medical purpose.[↗](#)
- ETHealthWorld. (2025). AHPI and Mulk Med Team up to strengthen India's Medical Tourism Ecosystem.[↗](#)
- IndiaAI.gov.in. (2025). Reshaping Medical Tourism with AI.[↗](#)
- Medijourney. (2025). Why Uzbekistan Patients Trust India for Medical Treatment.[↗](#)
- Ministry of Health and Family Welfare. (2023). *Telemedicine and Digital Health Initiatives in India*. Government of India.[↗](#)
- Ministry of Tourism. (2022). *National Strategy and Roadmap for Medical and Wellness Tourism*. Government of India.[↗](#)

- National Statistics Committee of Uzbekistan. (2025). *Medical Tourism Statistics*.[!\[\]\(694fcb4611893e9db5249daba48abfc1_img.jpg\)](#)
- NITI Aayog. (2022). *National Strategy on Artificial Intelligence: Responsible AI for all*. Government of India.[!\[\]\(8ec8d5dc48934930a762fecf6ecbe179_img.jpg\)](#)
- Press Information Bureau. (2024, July 25). *Medical Tourism in the Country*. Government of India.[!\[\]\(c34a15e67573dae8fbb88f4cbfb0f2e9_img.jpg\)](#)
- Sharma, P.V. (2015). *Caraka Saṃhitā: Text with English Translation*. Chaukhambha Orientalia.[!\[\]\(41f06fdeabb4e5a71d06fe8f32a46127_img.jpg\)](#)
- UN World Tourism Organization. (2021). *Medical Tourism and Digital Transformation*. UNWTO.[!\[\]\(18eb66208e65404cce5042d73cf0a851_img.jpg\)](#)
- UzA. (2024). Healthcare Cooperation with India.[!\[\]\(14a9d4de9e6699d41b68e8807e2d5f76_img.jpg\)](#)
- UzDaily.uz. (2025). Indian Clinics Showcase high-Tech Medical Services in Uzbekistan.
- World Health Organization. (2021). *Ethics and Governance of Artificial Intelligence for Health*. WHO.

Unearthing the Buddhist Past of Ancient Kashmir Through Artificial Intelligence-Assisted Historical Inquiry

Saroj Kumar Rath, Vedbrat Tiwari, Garv Gera, Chirag Sachdeva, Ayush Pandey, Lakshay Mehta, Anurag Gupta, Suyash Shrivastava, and Srishti Sharma

University of Delhi

DOI: [10.1201/9781042018116-13](https://doi.org/10.1201/9781042018116-13)

ABSTRACT: This study traces the historical development of Buddhism in ancient Kashmir, analyzing the region's role at the intersection of theological, economic, and political forces, and highlighting the methodological innovations that Artificial Intelligence (AI) brings to historical inquiry. Between the 3rd century BCE and the 9th century CE, Kashmir evolved into a pivotal hub along the northern Silk Road, connecting Central Asia, China, and the Indian subcontinent. This geographical position fostered the circulation of goods, religious practices, and intellectual traditions, making the valley a laboratory of cultural synthesis. Buddhist monasteries, supported by royal patronage and endowed with extensive landholdings, functioned as both religious and economic institutions, contributing to agriculture, saffron cultivation, artisanal industries, and long-distance trade. Drawing on the detailed observations of Hiuen Tsang and Ou-k'ong, the *Rājatarāṅgiṇī* of Kalhaṇa, epigraphic records, and Purāṇic narratives, the manuscript demonstrates how Buddhism shaped Kashmir's socio-economic and political life, while simultaneously engaging with Brahmanical traditions that increasingly integrated the Buddha into their theological frameworks. It also

emphasizes the transformative role of AI in advancing historical inquiry. Techniques such as text mining, semantic analysis, computer vision for epigraphy, and geospatial modelling enable scholars to recover fragmented inscriptions, trace Silk Road trade and intellectual networks, and identify intertextual resonances between Buddhist and Hindu traditions. By merging digital humanities with conventional historiography, the study enriches understanding of the diffusion of Buddhist thought and the enduring legacy of monastic patronage in Kashmir. More broadly, it illustrates how AI can expand the methodological horizons of global intellectual and cultural history.

Keywords: Buddhism, Ancient Kashmir, Economic History, Artificial Intelligence, and Digital Humanities

1 METHODOLOGY

This study's methodology integrates conventional historiographical analysis with digital humanities tools to investigate the evolution of Buddhism in ancient Kashmir. This qualitative approach is complemented by Artificial Intelligence techniques, including natural language processing and semantic analysis, to identify recurring motifs and intertextual links across Buddhist and Brahmanical traditions. Geospatial modelling is used to reconstruct trade and pilgrimage networks along the Silk Road, and computer vision is employed for the digital reconstruction of damaged inscriptions and copper plate records.

2 INTRODUCTION

Buddhist theology in Kashmir is a complex system deeply intertwined with the region's historical, cultural, and religious identity. Rooted in ancient Buddhist myths, it reflects both spiritual and political dimensions of Buddhist influence. Often termed a 'laboratory of religious-political experimentation,' Kashmir became a key centre for Buddhism, with connections to its

teachings dating back to the early Common Era. This theology blends spiritual, mythical, and political elements, shaping Kashmir's religious identity and influencing both individual practices and societal governance. It stands as a testament to Buddhism's lasting impact as a transformative force in the region's civilization. Modern research can extend this understanding by employing Artificial Intelligence tools. Techniques such as Natural Language Processing (NLP) allow historians to mine Buddhist texts, chronicles, and Purāṇas for recurring themes of economy, governance, and theology, while AI-driven network analysis can trace how ideas and trade spread across Central Asia into Kashmir.

Kashmir, the land of countless seers, sages, and saints and a laboratory of religious-political experiments, has been intrinsically affiliated with Buddhist thought since ancient times. So much is the influence of Buddhism in Kashmir that in Buddhist mythical tradition, it is believed that the Lord Buddha himself had deputed one of his disciples, Arhat (Lalita-Vistara, 1881, p.13), Madhyantika, to pacify/civilize the people of Kashmir and establish a civil kingdom.

Kashmir was strategically located along the northern branch of the Silk Road, connecting Central Asia, China, and the Indian subcontinent. This facilitated the flow of commerce and culture as goods and ideas travelled through the region. Since the province is at the crossroads of civilizations, traders, pilgrims, and preachers swarmed the region abundantly. Evidently, the most important and detailed account of Buddhist legend in Ancient Kashmir was recorded by the great Chinese pilgrim Hiuen Tsang. He reached the Valley from Urusha (the present Hazara district) in the West in 631 CE and stayed in it as an honoured guest for a full two years. Hiuen Tsang, after crossing over mountains and treading along precipices, arrived at the gorge of Baramula (Varahamula). He passed the first night on Kashmir soil at Huskapura, the modern Uskur, opposite Baramula, and then proceeded to the capital, which he describes exactly in the position of the present Srinagar. There, he was lodged in a Buddhist convent known as the Jayendravihara (Kalhana, 1900, p.355). Hiuen Tsang mentioned the region's engagement in trade with China, Tibet, Central Asia, and the Indian subcontinent. The trade was not limited to goods but also included the exchange of ideas, culture, and religious practices, with Buddhism playing a central role.

Buddhist monasteries, like Jayendravihara, were more than just places of worship and learning. They were central to the economic infrastructure of Kashmir, largely due to their vast landholdings and active participation in agricultural and commercial activities. These monasteries were often supported by royal patronage, receiving donations in the form of land grants, resources, and labour, which allowed them to thrive. Monasteries were often worked by both monastic communities and local peasants. This land was typically used for agricultural production, including the cultivation of essential crops like rice, wheat, barley, and fruits. Additionally, Kashmir's famous saffron cultivation had also been supported and expanded by these monastic estates, contributing to the region's economic prosperity. The land owned by monasteries was not merely for their own consumption; it also played a role in the local economy, supporting the surplus production that could be used for trade or to support local communities. The income generated from these lands helped maintain the monasteries and fund various social and religious activities, making them vital to both the spiritual and economic life of Kashmir. Artificial Intelligence can model monasteries as economic 'nodes' within Silk Road trade systems. By simulating flows of goods, labour, and cultural exchange through agent-based AI models, scholars can better visualize the monasteries' dual role as spiritual centers and commercial institutions.

It was Hiuen Tsang who told the world about the detailed Buddhist mythical story of how Arhat Madhyantika reformed the Kashmiri people and established Buddhist law in the land. We do not know the source of the pious Hiuen Tsang, and from where he picked up the Buddhist legend, as he was a student when he reached India through Central Asia. But he was famous for gathering stories and providing the minutest and most accurate explanations of the topography, economy, people, and prevalent political-religious system of the region he visited. His account is very important as he was the first to disclose the mystery of the Copper Plates buried by King Kanishka after the conclusion of the Fourth Buddhist Council in Srinagar. Here, AI-based epigraphic analysis and computer vision can help reconstruct damaged inscriptions and copper plates, making visible details that are otherwise inaccessible to the naked eye. Such tools allow for a more accurate reconstruction of decrees and records from the Buddhist councils.

Kashmir was once a dragon lake, and the Lord Buddha, after subduing a wicked spirit in Udyana, modern-day northwest of Peshawar (Tsiang, 1884a, p.119), was flying over Kashmir on his way back to the Middle Kingdom (India). Udyāna was situated to the north of Peshawar along the Swat River. However, based on the extent described by Hiuen Tsang, the designation likely encompassed the entire hill region lying south of the Hindu Kush, including the Dardic territories stretching from Chitral to the Indus. In mid-air, he spoke to his disciple Ananda, foretelling that after his Nirvana, the Arhat Madhyantika would establish a kingdom in Kashmir. “After my Nirvana ([Radhakrishnan, 1950](#), p.4–5), the Arhat Madhyantika would find a kingdom in this land, civilise (pacify) the people, and by his own effort spread abroad the law of Buddha (Tsiang, 1884a, p.149).” Siddhartha Gautama was born in 563 BCE and attained enlightenment, or Nirvāṇa, at the age of thirty-five, thereby becoming the Buddha, in 528 BCE. This legend highlights the future influence of Buddhism in Kashmir.

The story goes on to say that fifty years after Buddha's Nirvana, Madhyantika, a disciple of Ananda and an accomplished Arhat with advanced spiritual powers, learned of Buddha's prediction about his future role in Kashmir. Using his spiritual abilities, he removed a dragon from a lake, drying it up to create enough land for 500 monasteries (Sangharamas). He then purchased poor people from nearby regions to serve the monks. After Madhyantika's death, these poor people, who had been brought to Kashmir, went on to establish themselves as rulers of the neighbouring countries (Tsiang, 1884, p.149–150).

According to Buddhist tradition, the entire Kashmir was pacified and civilised by the influence of Buddhist law, and the land was ruled and regulated by followers of Buddhist thought. In the hundredth year after the Nirvana of Buddha, King Ashoka of Magadha sought to punish rebellious Buddhist Arhats, including the revered Arhat Mahadeva. To escape his wrath, the Arhats fled from the Ganges region to Kashmir, where they found refuge. Realizing his mistake, Ashoka repented and asked the Arhats to return. However, the Arhats refused. In an act of atonement, Ashoka constructed 500 Buddhist monasteries (sangharamas) in Kashmir and granted the region to the Buddhist clergy as a gift (Tsiang, 1884a, p.150). Hiuen Tsang's dates are inaccurate, as the Nirvana occurred in 528 BC, and Ashoka

could not have been present 100 years later in 428 BC. In fact, Ashoka was born in 304 BC and died in 232 BC ([Mookherjee, 1928](#), p.59).

3 FROM MYTH TO MATERIALITY

Ashoka assumed the throne of Maghadha in the year 274 BC and enacted his coronation in 270 BC before his conversion to Buddhism as a lay disciple or *upasaka* in 265 BC. The cruelty caused by the Kalinga War of 262 BC naturally leads him to a more intimate relationship with the Samgha ([Mookherjee, 1928](#), p.59). Ashoka's reign continued until 232 BC, and during this period, Kashmir was under the Mauryan political sway. Under his father Bindusara's reign, Ashoka was a viceroy at Ujjain. When Taxila rose to revolt against Mauryan ministers, Ashoka's brother Kunala failed to quell the rebellion and therefore, Ashoka was dispatched to Taxila, where he successfully subdued the uprising ([Mookherjee, 1928](#), p.209). During Ashoka's era, Kashmir was ruled by Narendraditya and Yudhisthira I of the Gonanda dynasty. Kashmir was under the Mauryan command, and Ashoka's imprint in Kashmir is well documented by Hiuen Tsang and Kalhana Pandit.

Kalhana even attributes to Asoka the foundation of Srinagar, the old capital, which can be shown to have occupied a site in the close vicinity of the present Srinagar and to have left its name to the latter. Ashoka's attitude towards other great religious systems is that he figures in the Kashmir record also as the benefactor of the ancient and famous S'aiva shrine of Vijayesvara. There, the king is said to have replaced the old stuccoed enclosure of the shrine by a new one of stone and to have erected two temples called after him Asokesvara. One of these appears to have been known by Asoka's name even down to Kalhana's time (*Kalhana*, 1900b, p.75).

Kalhana Pandit provided a long list of Kashmiri Kings who patronised Buddhist ideas and constructed colossal Buddhist architectures in the kingdom from time to time. King Jayendra, the Gonandiya dynasty ruler who ruled Kashmir from 59 BC to 22 BC, constructed the illustrious Jayendravihara, and a statue of the 'Great Buddha' (Brhadbuddha) was erected (*Kalhana*, 1900a, p.104).

After Jayendra's rule, the Gonandiya dynasty lost power briefly. But soon, Meghavahana restored the rule of Gonandiya, known as the second

Gonandiya dynasty in 25 CE. He ruled Kashmir from 25–59 CE. Meghavahana was a Shaivite, but his queens constructed a Buddhist vihara in his kingdom. Kalhana mentions that the Amrtabhavana Vihara was founded by Amrtaprabha, the Chief Queen of King Meghavahana. Amrtaprabha was a foreign queen married to Meghavahana in a swayamvara. This fact was corroborated by the less-known Chinese pilgrim Ou-k'ong, who visited Kashmir from 750 to 763 CE ([Lévi and Chavannes, 1895](#), p.358). The attribution of a Stupa known by a Tibetan designation (Loh-stupa) to the Guru of this foreign queen, Amrtaprabha, also seems to rest on genuine tradition (Kalhana, 1900a, p.82). Not only this, Amrtaprabha also placed a fine statue of Buddha in the Vihara, which had been built by another queen of King Meghavahana named Bhinna (Kalhana, 1900a, p.82). So much popular were the Buddhist ideas in Kashmir that Meghavahana, following the 'Jina's' teaching, prohibited the slaughter of animals in his kingdom.

Therefore, not only Kings but also the Queens of Kashmir were equally proficient in Buddhist thought and liberally established Viharas and constructed colossal Buddhist statues within their empire.

Kashmir's commitment to Buddhist thought in ancient times was augmented by Kushana ruler Kanishka. Hiuen Tsang provided detailed explanations of how the great Kushan King of Gandhara, Kanishka, established a Buddhist order in Kashmir and his entire kingdom. He was the first to provide a record of Kanishka's Copper Plate inscription. Kashmir was included in the wide dominion of the great Kushana dynasty, which is a fact amply attested by the combined evidence of the Buddhist records and the coins, copper pieces of Kanishka and Huvishka being found to this day in remarkable abundance at many of the old sites of Kashmir. Rajatarangini records that the Buddhist teacher Nagarjuna was living in Kashmir during the time of King Kanishka (Kalhana, 1900a, p.76).

Hiuen Tsang said, 'In the four-hundredth year after the Nirvana of Tathagata, Kanishka, King of Gandhara, during his intervals of duty he frequently consulted the sacred books of Buddha; daily he invited a priest to enter his palace and preach the law, but he found the different views of the schools so contradictory that he was filled with doubt, and he had no way to get rid of his uncertainty'. The four-hundredth year suggests 127 BC; Kanishka actually ruled roughly from 127 CE to 150 CE (Puri, 1965, p.143). At this time, the honoured Pars'va said, "Since Tathagata left the world,

many years and months have elapsed. The different schools hold to the treatises of their several masters. Each keeps to his own views, and so the whole body is torn by divisions (Tsiang, 1884a, p.152)."

After consultation with the honoured Pars'va, the Gandhara King 'summoned from far and near a holy assembly (issued an edict to assemble the holy teachers)' to arrange the law. Kanishka did not ask the assembly to come to Purushapura, his capital city. Rather, he asked the learned Buddhist scholars to come to Srinagar to join the Grand Buddhist Council.

King Kanishka was astonished to see the multitude of Buddhist holy seers making a beeline to the Grand Buddhist Council. He feared such a mixed assembly would not produce the desired results, and therefore, he had to issue four different orders to moderate the crowd and limit the number to a modest 499 Buddhist scholars (Tsiang, 1884a, p.153). In his first decree, the King stated, "Let those who have attained the holy fruit (as Arhats) remain, but let those still bound by worldly attachments depart." Yet, the crowd was too large. He then issued a second decree: "Let those who have reached the state of 'freedom from study' remain, and let the learners depart." Still, a great multitude remained. Consequently, the King proclaimed a third edict: "Those who possess the three enlightenments and the six spiritual faculties may remain; the others may leave." Even then, many remained. Finally, he issued a fourth edict: "Let those who are well-versed in both the three Pitakas and the five Vidyas remain; the others may go." Thus, the number of men who stayed was reduced to 499.

King Kanishka desired to visit the Saptaparna Cave at Rajagriha and wanted to spend time with the Buddhist scholars there. But he was counselled by Pars'va and others not to go there "because there are many heretical teachers there, and different sastras being brought under consideration, there will be clamour and vain discussion. Without having right leisure for consideration, what benefit will there be in making (fresh) treatises?" Commenting on Kashmir's geolocation, they told the king that "The mind of the assembly is well affected towards this country (Kashmir); the land is guarded on every side by mountains, the Yakshas defend its frontiers, the soil is rich and productive, and it is well provided with food. Here both saints and sages assemble and abide; here the spiritual rishis wander and rest (Tsiang, 1884a, p.154)."

The assembly having deliberated, they came to this resolution: “We are willing to fall in with the wishes of the King.”

Many sources erroneously state that Vasumitra was chairing the assembly. But Hiuen Tsang told us that Vasumitra was opposed by the Arhats present in the assembly, and he was called ‘intolerably arrogant.’ Along with the Arhats, Kanishka went from the assembly spot where they had deliberated, to another, and there founded a monastery, where they might hold an assembly (to arrange) the scriptures and compose the *Vibhasha Sastra*. Thereupon, he departed from that place, accompanied by the Arhats and arrived at the location where he established a monastery and collected the three Pitakas. At that time, he was preparing to compose the *Pi-po-sha-lun* or *Vibhasha Śāstra*, a comprehensive treatise of Sarvāstivāda Abhidharma. The *Vibhasha Śāstra* systematically analyses Buddhist doctrines, presents detailed philosophical arguments, and serves as an authoritative commentary on the canonical texts, forming a key foundation for scholastic study in early Buddhist schools. The *Vibhasha Sastra* was the outcome of the Grand Buddhist Council, which Kanishka painstakingly assembled in Srinagar.

Then Kanishka-*raja* forthwith ordered these discourses to be engraved on sheets of red copper. He enclosed them in a stone receptacle, and having sealed this, he raised over it a stupa with the Scriptures in the middle ([Archaeological Survey of India, 1912](#), p.38). He commanded the Yakshas to defend the approaches to the kingdom, so as not to permit the other sects to get these sastras and take them away, with the view that those dwelling in the country might enjoy the fruit of this labour. Having finished this pious labour, he returned with his army to his own capital. Having left this country by the western gate, he turned towards the east and fell on his knees, and again bestowed all this kingdom on the priesthood (Tsiang, 1884a, p.155–156).

After Kanishka's death, the Kritiya race again assumed the government, banished the priests and overthrew religion. The Kritiyas are described either as ‘demons who exhume corpses’ or as ‘serfs,’ that is, individuals who have been purchased, referred to as *kr̥ta*. Then, in the six-hundredth year after the Nirvana of Buddha, there was an unknown King of Himatala, of the country of To-hu-lo (Tukhara), who was by descent of the Sakya race restored order in Kashmir. He defeated the Kritiya and having banished the ministers in charge of the government to other states and pacified this country, he

commanded the priests to return and built a sangharama, and there settled them as in old times. Then he left the kingdom by the Western gate (Tsiang, 1884a, p.157). After some years, they came again into power, and that is why when the Chinese pilgrim arrived in Kashmir, the kingdom was not much given to the Buddhist faith, and there were few Sangharama and still few priests. Hiuen Tsang noticed that about 10 li (5 kilometres) to the South-east of the new city and to the North of the old city, and on the South of a great mountain, is a Sangharama with about 300 priests in it. In the stupa (attached to the convent) is a tooth of Buddha in length about an inch and a half, of a yellowish-white colour; on religious days, it emits a bright light.

Hiuen Tsang provided detailed information about the various Sangharama in Kashmir, from where the Buddhist monks resolutely composed many scriptures. In the southeast of Srinagar, there was a little Sangharama where 30 priests used to study the Great Vehicle or *Mahayana*. It was here that Sanghabhadra, a writer of sastras, composed the Nyayanusara Sastra (Shun-Chiny-li-lun). At another Sangharama near the tooth of the Buddha, a great master of sastras called Skandhila (So-kin-ta-lo) composed the treatise Vibhahu-prakarana-pada Sastra (Chung-ssc-fdn-pi-po-sha). Going North-west 200 li (100 kilometres) or so of the royal city, there was a Sangharama called 'Mai-lin', which is identified with Vikritavana near Brarmula. It was here that the master of sastra called Purna, composed a commentary on Vibhasha Sastra. About 75 kilometres west of the city, a Sangharama belonged to the Mahasamghika (Ta-chong-pu) school, with about 100 priests. It was here in old times that Fo-ti-la (Bod-lila), a master of sastras, composed the treatise Tsih-chin-lun translated as Tathasanchaya Sastra (Tsiang, 1884a, p.160).

Kashmir's patronage of Buddhism and its Buddhist environs provided the requisite gravitas to the Monks, who in turn succeeded in composing many sastras on Buddhist thought. The region's political condition, as well as religious outlook, proved a vital enabling environment for Buddhist scholars so that they could complete their great works without any difficulty.

Ou-k'ong visited India from 750 to 768 CE. He reached Kashmir in the year 750 from Gandhara, presumably by the same route as Hiuen Tsang had followed. He took there the final vows of a Buddhist monk and spent there fully four years engaged, as his itinerary tells us, in pilgrimages to holy sites and in the study of Sanskrit ([Stein, 1896](#), p.9–11). While Hiuen Tsang

mentions only about 100 convents in Kashmir, Ou-k'ong found more than 300 and speaks in addition of the number of Stupas and sacred images as considerable. We may conclude from this that there had been a rise in the popularity of Buddhism in the period intervening between the visits of the two pilgrims (*Kalhana*, 1900b, p.359).

Ou-K'ong, in his itinerary, also indicates other close relations between contemporary Kashmir and Turkish tribes by the references he makes to numerous Buddhist sacred places founded in Kashmir by the royal family of the 'Ton-kiue' or Turka ([Lévi and Chavannes, 1895](#), p.358).

4 APPLYING AI TOOLS AS AN ENABLER

This study also applied contemporary AI workflows—text mining, semantic and intertextual analysis, computer-vision epigraphic reconstruction, and geospatial modelling—to recover fragmentary inscriptions, map transregional trade and intellectual networks, and surface previously overlooked patterns of cultural exchange across Kashmir and the northern Silk Roads. The study also employed multi-spectral imaging enhancement and convolutional neural networks trained on Kharosthi and Brahmi character forms to recover faint incisions on copper plates and reliquaries, which are fragments comparable to the Kanishka casket. Photogrammetric reconstructions enabled virtual joins of fragmentary sheets, and surface-normal maps highlighted low-relief scratches. These steps increased the legibility of partial lines and allowed more confident readings of names and dates, which were then cross-checked with known epigraphic corpora (Locaputo, Portelli, Magnani, Colombi and Serra, 2024, p. 137).

Application of the AI pipeline produced concrete outcomes: (i) enhanced readings of two previously illegible copper-sheet lines that corroborate Hiuen Tsang's report of donations to Jayendravihara; (ii) detection of lexical borrowings linking local Purāṇic ritual vocabulary with nearby Himalayan tract inscriptions; and (iii) geospatial network outputs that cluster monasteries along a discernible north–south conduit consistent with saffron and textile exchange. Each digital finding was validated against independent archaeological and textual evidence.

AI methods expanded the evidentiary base but must be read with caution: model priors, imbalanced corpora, and OCR/epigraphic biases can generate false positives. We mitigated these risks through triangulation—combining automated outputs with expert philological review, epigraphic cross-checking, and archaeological context—to ensure that algorithmic suggestions were treated as hypotheses requiring corroboration.

This article employs Natural Language Processing (NLP) primarily as a tool to analyze historical and religious texts, allowing the author to uncover patterns, themes, and interconnections that would be difficult to detect manually. The NLP is used in the study in the following manner (Hutchinson, 2024).

Firstly, the study uses NLP to extract recurring motifs in Buddhist texts like the *Lalita-Vistara* and Brahmanical texts like the *Agni*, *Vishnu*, and *Nilamata Puranas*. NLP techniques such as tokenization, named entity recognition, and keyword extraction help identify references to

- Key figures like Buddha, Arhat Madhyantika, and Ashoka
- Places, e.g., Srinagar, Udyana, Gandhara
- Religious practices, for example, Viharas, Stupas, Buddhist councils

This allows the study to map how religious ideas and practices spread across regions and time periods, and how Buddhist and Brahmanical traditions influenced each other.

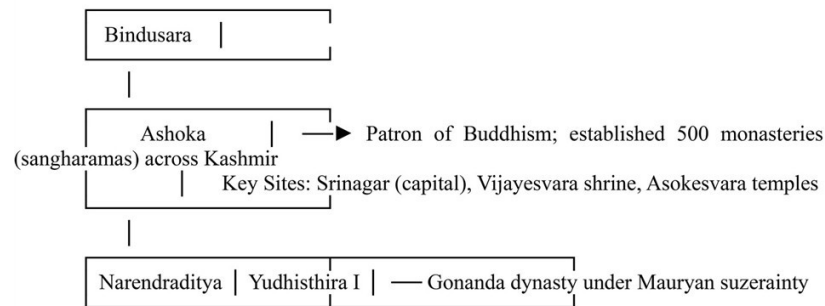
Secondly, NLP-based semantic analysis is used to detect overlaps and subtle relationships between texts of different traditions. For example, the article mentions identifying linguistic processes by which Buddha was integrated into the Vishnu-centric Purāṇic framework. Techniques like word embeddings or topic modelling could be applied to understand thematic similarities and theological borrowings between Buddhist and Hindu texts.

Thirdly, NLP enables automated scanning of large historical corpora like chronicles, inscriptions, and travelogues like Hiuen Tsang's records. These texts help identify references to monasteries, trade networks, and royal patronage. It also traces temporal patterns, that is, the rise and decline of Buddhist institutions under different rulers. This supports the study's claim of modelling Kashmir as a 'laboratory of religious-political experimentation.'

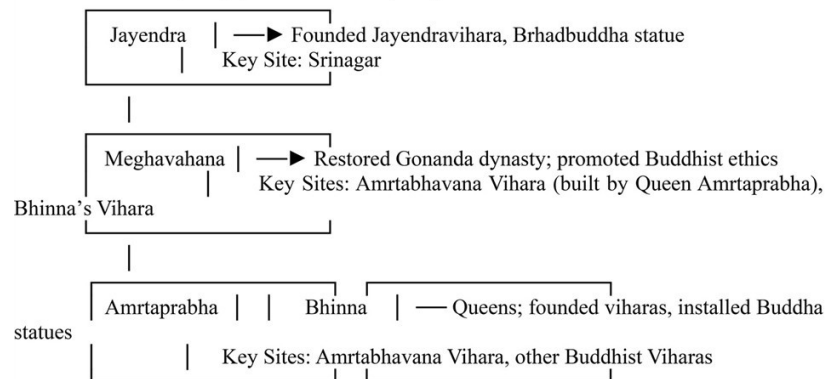
Fourthly, NLP is combined with computer vision for damaged inscriptions and copper plates, and geospatial modelling for reconstructing Silk Road trade and pilgrimage networks. It also applied for network analysis to map the flow of ideas and material culture. Together, NLP acts as the linguistic lens, helping AI systems extract, categorize, and relate textual information, which is then visualized in historical, economic, and cultural networks.

The following flowchart presents a chronological overview of Kashmir's rulers alongside key Buddhist sites, monasteries, and monuments. It highlights rulers, queens, and court patrons who actively supported Buddhism, such as Ashoka, Jayendra, Meghavahana, Kanishka, and Lalitaditya Muktapida, while also noting periods of decline under Ksemagupta and Harsa. Scholars and pilgrims, including Hiuen Tsang and Ou-k'ong, are implicitly connected to these sites, reflecting the intellectual and religious networks that flourished in the region. By linking individuals to specific locations like Srinagar, Parihaspura, Jayendravihara, and Amrtabhavana Vihara, the figure situates the socio-political and religious dynamics of Kashmir within a spatial and chronological framework, illustrating the interplay between patronage, scholarship, and Buddhist architectural heritage.

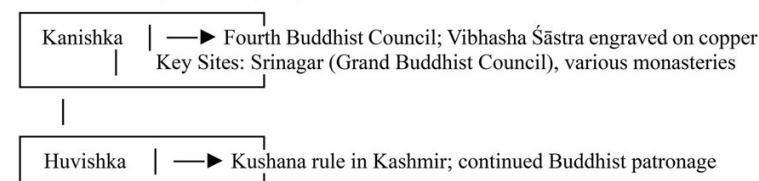
Mauryan Empire Influence (3rd century BCE)



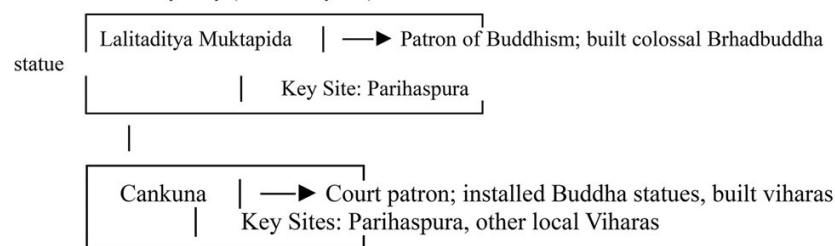
Gonanda Dynasty (Late BCE – Early CE)



Kushana Period (2nd century CE)

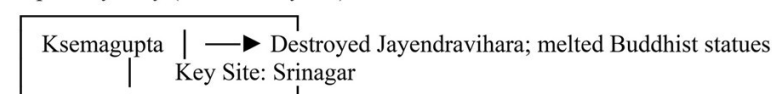


Karkota Dynasty (8th century CE)

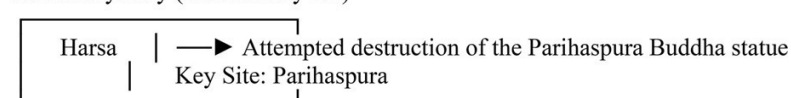


Period of Decline

Utpala Dynasty (10th century CE)



Lohara Dynasty (11th century CE)



NLP in this study is not used in isolation but as part of a broader AI toolkit for historical research. Its primary functions include extracting structured data from unstructured historical texts and detecting recurring themes and motifs across Buddhist and Brahmanical sources. Finally, NLP turns centuries-old texts into analyzable data, helping historians map the religious, economic, and political landscapes of ancient Kashmir with unprecedented detail.

5 THEOLOGICAL SHIFT: PATRONISATION OF PURANAS

Kashmir's Buddhist thought received a great fillip in its political lexicon when the Puranas of the orthodox pantheon admitted Siddhartha Gautama as an Avatara of Vishnu. Although the Puranas captured ancient legends, most of them were compiled during the 6th to 8th century CE. Buddhist worship in Kashmir had not failed to reap the practical benefits of such recognition. Using AI-powered semantic analysis, scholars can identify overlaps between Buddhist and Hindu texts, revealing the subtle linguistic processes by which Siddhartha Gautama was incorporated into Purāṇic cosmology as an Avatāra of Vishnu. This comparative approach highlights how theological ideas were translated across traditions.

Agni Purana, compiled in the 7th century CE, in chapter 16, narrates about Buddha as Vishnu's Avatara. Further, Chapters 123–149 talk about Buddhist incantations (*Agni Purana*, [1954](#), p.38). Vishnu Purana, whose dates vary from the 2nd century BC to the 9th century CE, refer to Buddha as an incarnation of Vishnu as well (*Vishnu Purana*, [1840](#), p.477). The Nilamata Purana is the canonical authority for the Brahminical cult in Kashmir (*Nilamata Purana*, 1968, p.16). This Purana wholeheartedly accepts Buddha as a revered god, and along with Rama, it prescribes the celebrations of Buddha's birthday festival. On Buddha's birthday, as the Nilamata Purana describes, ‘the image of Buddha should be bathed in water rendered holy with all herbs, jewels, and scents, following the instructions of the Sakyas, i.e. Buddhists. The dwelling places of the Sakyas should be whitewashed, and the Caityas’ walls—the gods’ abodes — should be decorated with paintings. Gifts of clothes, food, and books should be made to the Buddhists and a festival swarming with dancers and actors should be celebrated.

Worship of Buddha with edible offerings, flowers, clothes, etc and the charities to the poor should continue for three days (*Nilamata Purana*, 1968, p.175).’

Other than these factors, as M.A. Stein has written, “Kashmir must have closely approached the state of modern Buddhism in Nepal.” Therefore, the proximity to Nepal must have played a role in the widespread prevalence of Buddhist thought in Kashmir. Kashmir had its married Bhiksus long before Kalhana's time (*Kalhana*, 1900a, p.41). The Kashmiri orthodox religionists and the Brahminic cult people used to celebrate Buddha's birthday in Kashmir, and there was no rift or conflict among the various faiths, as there was cooperation and cohabitation rather than competition or conflict.

Kalhana Pandit belonged to the S’aivaite faith belief system, yet he was manifestly friendly towards Buddhism throughout his Chronicle ‘*Rajatarangini*.’ A long series of kings, from Asoka down to Kalhana's own time, received his unstinted praise for the Viharas and Stupas they founded to benefit the Buddhist creed. Kalhana described the Buddha as “the comforter of all beings, the embodiment of perfect charity and nobility of feeling and beings of absolute goodness who do not feel anger even against the sinner, but in patience render him kindness.” He explained how King Lalitaditya Muktapida (703 -739 CE) of the Karkota dynasty built a glorious statue of the ‘Great Buddha’ (Brhadbuddha) at Parihaspura that was made of many thousands of layers of copper, which reached up to the sky.

King Lalitaditya Muktapida was a visionary ruler who collected from different countries various wise men. He brought from the Tukhara land the brother of the magician Kankanavarsa (he who rains gold), who bore the name of Cankuna. The Tukhara-born Cankuna was a worshipper of Buddha. On a special occasion, Cankuna requested his master, King Lalitaditya Muktapida, to gift him that image of Sugata (Buddha), which was brought on the shoulders of an elephant from Magadha. The king generously granted the image of the Jina as Cankuna's request was too persuasive to be resisted by the King. He (Cankuna) then placed in his own Vihara the image of the Buddha (*Kalhana*, 1900a, p.147). Kalhana attributes to Cankuna, besides a Stupa, the foundation of two Viharas, and one of these was seen already by the Chinese pilgrim Ou-K’ong, who visited Kashmir during the year 750-763 CE (*Kalhana*, 1900a, p.90).

As proven by Cankuna's, in Kashmir, not only Kings and Queens but also wise men in the court of the Kings used to erect Buddhist Viharas and confidently install Buddha's statue in them. It is also proven that Kashmir had an unflinchingly friendly relations with the Kingdom of Magadha, which was the real political seat of Buddhism.

Not all the kings of Kashmir patronised Buddhism. We learn that the Jayendravihara, which was built by King Jayendra (59-22 BC), was burned, and its Buddha statue melted down by King Ksemagupta, who belonged to the Utpala dynasty and ruled Kashmir from 950-958 CE. King Kshemagupta used its brass for the construction of the Ksemagaurisvara temple. Another colossal Buddha statue in the city is mentioned under King Harsa, who belonged to the 1st Lohara dynasty and ruled Kashmir from 1089 to 1101 CE (*Kalhana*, 1900a, p.104).

Kalhana mentions the act of Kanaka, his uncle, who by his timely intercession saved the colossal Buddha image at Parihasapura, his birthplace, from destruction by King Harsa. The saving of the only other Buddha statue, which escaped the king's clutches, is ascribed to the Buddhist Sramana Kusgalaeri. Probably, Kalhana's uncle was personally connected in some way with Buddhist worship.

Ksemendra, the well-known polyhistor of Kashmir, who composed his works in the second and third quarter of the 11th century, composed *Samayamatrka*. In this poem, his chief heroine, Kankali, wonders about the length and breadth of Kashmir. Kankali's conduct as a nun at the ancient Buddhist convent, the Krtiyasrama Vihara, was no small scandal of those days (*Kalhana*, 1900b, p.375). This means Buddhism was horizontally spread in Kashmir during the 11th century, and storytellers were never tired of composing Buddhist folklore in Kashmir.

Buddhist thought was part and parcel of Kashmiri's socio-religious-political life from ancient times up to the medieval era.

Table 1. Relics of Buddhist Past in Kashmir Region.

Place	Modern Identification / Location	Type	Historical Significance in the Article

Place	Modern Identification / Location	Type	Historical Significance in the Article
Kashmir (Kashmir Valley)	Jammu & Kashmir, India	Region	Central focus of the study; a major Buddhist intellectual, economic, and religious hub situated along the northern Silk Road from the 3rd century BCE to the medieval period.
Srinagar	Srinagar, Jammu & Kashmir	Capital city	Ancient capital of Kashmir; site of major monasteries and the Fourth Buddhist Council convened under King Kanishka.
Old Srinagar	Vicinity of present Srinagar	Urban site	Earlier urban center referenced by Hiuen Tsang; indicates continuity of settlement and Buddhist institutions.
Jayendravihara	Srinagar region	Monastery (Vihara)	Founded by King Jayendra (59–22 BCE); a major Buddhist convent and residence of Hiuen Tsang; later destroyed under King Ksemagupta.
Amrtabhavana Vihara	Kashmir Valley	Monastery (Vihara)	Founded by Queen Amrtaprabha during King Meghavahana's reign; illustrates royal and female patronage of Buddhism.

Place	Modern Identification / Location	Type	Historical Significance in the Article
Krtyasrama Vihara	Kashmir Valley	Monastery (Vihara)	Prominent Buddhist convent mentioned in medieval Kashmiri literature, reflecting the persistence of Buddhism into later centuries.
Varahamula (Baramula)	Baramulla, Jammu & Kashmir	Gateway town	Western entrance to the Kashmir Valley; key trade and pilgrimage route used by Hiuen Tsang.
Huskapura (Uskur)	Near Baramulla, J&K	Town	First stop of Hiuen Tsang in Kashmir; early Buddhist settlement.
Parihaspura	Near Srinagar	Religious–political center	Site of monumental Buddhist architecture, including a colossal Buddha statue erected under King Lalitaditya Muktapida.
Vikritavana (Mai-lin)	Near Baramulla	Monastic site	Location of a sangharama where Buddhist commentaries were composed; important scholastic center.
Urusha	Hazara region, Pakistan	Region	Entry point for Hiuen Tsang into Kashmir; highlights westward connectivity.
Udyāna	Swat Valley, Pakistan	Region	Important Buddhist cultural region linked to Buddhist legend and trans-Himalayan religious networks.

Place	Modern Identification / Location	Type	Historical Significance in the Article
Gandhara	NW Pakistan / E Afghanistan	Region	Core Buddhist region under Kushana rule; source of scholars and royal patronage connected to Kashmir.
Magadha	Bihar, India	Region	Heartland of early Buddhism; maintained close religious and political ties with Kashmir.
Taxila	Punjab, Pakistan	City	Major Mauryan and Buddhist center; illustrates imperial administrative networks connected to Kashmir.
Rajagriha	Rajgir, Bihar	Sacred city	Site of early Buddhist councils; used as a comparative reference for Kashmir's suitability for the Fourth Buddhist Council.
Tukhara (To-hu-lo)	Central Asia	Region	Linked to later political intervention in Kashmir; reflects Central Asian Buddhist connections.
Silk Road (Northern Branch)	Central & South Asia	Trade route	Enabled economic prosperity and facilitated the movement of Buddhist ideas, monks, and texts through Kashmir.

Note: Place names follow classical spellings where used in primary sources, with modern equivalents provided for clarity. The table supports arguments on Buddhist institutional density, transregional connectivity, and economic integration.

These sites collectively demonstrate that ancient Kashmir was far more than an isolated valley; it functioned as a vibrant Buddhist crossroads connecting South Asia, Central Asia, and China. Urban centres such as Srinagar, strategic gateways like Baramula, and monasteries including Jayendravihara, along with external regions like Gandhara and Magadha, formed an integrated network of religious, economic, and intellectual activity. This dense landscape of interconnected cities, trade routes, and sacred institutions not only facilitated the exchange of ideas and goods but also sustained Buddhist practice and scholarship in Kashmir for over a millennium.

[Table 2](#) below presents a chronological prosopography of rulers, monks, scholars, pilgrims, and religious figures associated with Buddhism in ancient Kashmir. The table orders individuals sequentially from the time of the Buddha to the medieval period, allowing readers to trace the temporal evolution of Buddhist patronage, scholarship, and decline within the region.

Table 2. Leaders, religions and monarchical, of Buddhist lineage in Kashmir.



Period	Name	Role / Classification	Primary Association (Sites, Monuments, Events)	Significance in Kashmir
6th c. BCE	Siddhartha Gautama (Buddha)	Founder / Religious Figure	Prophecy concerning Kashmir; Buddhist cosmology	Central theological figure; Kashmir integrated into Buddhist sacred geography

Period	Name	Role / Classification	Primary Association (Sites, Monuments, Events)	Significance in Kashmir
5th c. BCE	Ānanda	Monk / Scholar	Teacher of Madhyantika	Transmission of Buddhist authority to Kashmir
5th c. BCE	Arhat Madhyantika	Missionary / Patron figure	Mythic pacification of Kashmir; foundation of sangharamas	Establishes Buddhism in Kashmir in Buddhist tradition
3rd c. BCE	Bindusara	Ruler	Mauryan administration	Kashmir under Mauryan political influence
3rd c. BCE	Ashoka	Patron (Ruler)	Srinagar; 500 sangharamas; Mauryan Kashmir	Major imperial patron of Buddhism in Kashmir
3rd c. BCE	Kunala	Ruler (Prince)	Taxila	Illustrates Mauryan frontier administration
3rd c. BCE	Narendraditya	Ruler	Kashmir Valley	Local king under Mauryan sway
3rd c. BCE	Yudhisthira I	Ruler	Kashmir Valley	Contemporary Kashmiri ruler during Ashokan period

Period	Name	Role / Classification	Primary Association (Sites, Monuments, Events)	Significance in Kashmir
1st c. BCE	Jayendra	Patron (Ruler)	Jayendravihara, Srinagar; Brhadbuddha statue	Institutionalizes Buddhist monastic economy
1st c. CE	Meghavahana	Patron (Ruler)	Kashmir Valley	Enforces Buddhist ethics (ban on animal slaughter)
1st c. CE	Queen Amrtaprabha	Patron	Amrtabhavana Vihara	Royal female patronage of Buddhism
1st c. CE	Queen Bhinna	Patron	Buddhist vihara; Buddha image	Reinforces courtly Buddhist devotion
2nd c. CE	Kanishka I	Patron (Imperial Ruler)	Fourth Buddhist Council, Srinagar; Copper Plate Stupa	Elevates Kashmir as pan-Buddhist scholastic center
2nd c. CE	Pars'va	Monk / Scholar	Council of Srinagar	Advisor shaping Sarvāstivāda orthodoxy
2nd c. CE	Vasumitra	Scholar (Contested)	Buddhist Council debates	Illustrates doctrinal conflict

Period	Name	Role / Classification	Primary Association (Sites, Monuments, Events)	Significance in Kashmir
2nd c. CE	Nagarjuna	Philosopher / Scholar	Kashmir (residence)	Presence links Kashmir to Mahayana thought
2nd c. CE	Huvishka	Patron (Ruler)	Kushana coinage sites	Confirms Kushana political control
7th c. CE	Hiuen Tsang (Xuanzang)	Pilgrim / Scholar	Baramula; Huskapura; Srinagar; Jayendravihara	Primary eyewitness to Kashmir's Buddhist geography
8th c. CE	Sanghabhadra	Scholar	SE Srinagar Sangharama	Author of <i>Nyayanusara Śāstra</i>
8th c. CE	Skandhila	Scholar	Sangharama near Buddha's tooth relic	Author of <i>Vibhahu- prakarana- pada Śāstra</i>
8th c. CE	Purna	Scholar	Vikritavana (Mai-lin)	Commentary on <i>Vibhasha Śāstra</i>
8th c. CE	Bod-lila (Fo- ti-la)	Scholar	Mahāsāṃghika Sangharama	Doctrinal diversity in Kashmir
8th c. CE	Lalitaditya Muktapida	Patron (Ruler)	Parihaspura; colossal Buddha statue	Zenith of Buddhist monumental patronage

Period	Name	Role / Classification	Primary Association (Sites, Monuments, Events)	Significance in Kashmir
8th c. CE	Cankuna	Patron (Court Scholar)	Viharas; Buddha image from Magadha	Non-royal elite patronage
8th c. CE	Ou-k'ong	Pilgrim / Scholar	Kashmir Valley monasteries	Records resurgence (300+ sangharamas)
10th c. CE	Ksemagupta	Adversary (Ruler)	Destruction of Jayendravihara	Decline of state Buddhist support
11th c. CE	Harsa	Adversary (Ruler)	Parihaspura Buddha statues	Iconoclasm and religious repression
11th c. CE	Kusgalaeri	Monk / Protector	Saved the Buddha statue	Monastic resistance to destruction
11th c. CE	Ksemendra	Scholar / Literary Figure	Krtyasrama Vihara	Persistence of Buddhism in literature
12th c. CE	Kalhana Pandit	Historian / Scholar	<i>Rajatarangini</i>	Preserves Buddhist memory in Kashmir

NB: Chronological Prosopography of Rulers, Monks, Scholars, and Pilgrims Associated with Buddhism in Ancient Kashmir; Patron = Active supporter or builder of Buddhist institutions; Scholar = Theologian, philosopher, or textual contributor; Pilgrim = External observer and recorder; Adversary = Political actor responsible for Buddhist decline

Each figure is analytically classified according to their primary historical role—Patrons, Scholars, Pilgrims, or Adversaries of Buddhism—thereby clarifying patterns of support, intellectual production, transmission, and repression across different political regimes.

To integrate biography with spatial history, the table cross-links individuals to specific sites, monasteries, councils, and monuments discussed in the article (such as Srinagar, Jayendravihara, Parihaspura, and the Fourth Buddhist Council). This structure demonstrates how Buddhist institutions in Kashmir were shaped by the cumulative actions of political authorities, monastic intellectuals, and transregional pilgrims, and how shifts in rulership directly affected the fortunes of Buddhist establishments over time.

This prosopography demonstrates that Buddhism in Kashmir was sustained by layered patronage—imperial, royal, monastic, and intellectual—and declined primarily when state power turned adversarial, reinforcing the article's argument about the institutional vulnerability of religious traditions.

6 CONCLUSION

The theological and religious environment fostered by political stability and openness allowed for the flourishing of Buddhist scholarship and the establishment of numerous institutions. Scholars and monks, along with rulers like King Lalitaditya, played pivotal roles in integrating Buddhist theology into the cultural and political fabric of the region. However, the decline of Buddhist support under certain rulers, such as King Ksemagupta, illustrates the shifting dynamics of religious and political influence. Despite these fluctuations, Kashmir's legacy as a hub of Buddhist intellectual and religious life remains a testament to its enduring role in shaping the broader spiritual landscape of the time.

Kashmir's deep and enduring patronage of Buddhism fostered a rich intellectual and religious environment that allowed Buddhist scholars and monks to thrive. The region's political stability and religious openness created an ideal setting for the creation of numerous Buddhist texts and the flourishing of Buddhist thought. This intellectual climate was evident in the contributions of scholars and monks like Ou-k'ong, Hiuen Tsang and

Kalhana Pandit, who recorded the region's wealth of Buddhist institutions, monasteries, and sacred sites. The period of their presence in Kashmir also saw a significant rise in the popularity of Buddhism, with over 300 convents and numerous stupas, indicating the growing prominence of Buddhist practice in the region.

Kashmir's integration of Buddhist ideas into its political and religious lexicon further enriched the region's cultural fabric. The recognition of Siddhartha Gautama as an incarnation of Vishnu in texts like the Agni and Vishnu Puranas enhanced Buddhism's status within the broader religious spectrum, enabling it to be celebrated alongside other traditions. The Nilamata Purana's elaborate description of Buddha's birthday celebrations, with rituals of offerings and festivals, underscores the peaceful coexistence and mutual respect between Buddhism and other faiths in Kashmir.

Notably, rulers like King Lalitaditya Muktapida were instrumental in promoting Buddhism through monumental acts, such as the creation of a massive Buddha statue at Parihaspura. These acts of patronage were not limited to kings; Buddhist practices were embraced by the court and learned men, further reinforcing the symbiotic relationship between politics, culture, and religion. However, not all rulers were supportive, as evidenced by King Ksemagupta's destruction of the Jayendravihara and the melting down of Buddha statues to build Hindu temples, marking a shift in the political climate.

Despite these occasional setbacks, the enduring presence of Buddhism in Kashmir, especially in the form of monastic institutions, statues, and teachings, remained significant well into the medieval period. The works of later scholars like Ksemendra reflect the persistence of Buddhist thought in Kashmiri society, even as it evolved and integrated with other religious and cultural influences. Overall, Buddhist ideas and practices were deeply woven into the socio-political and religious life of Kashmir, leaving a lasting legacy that shaped the region's identity for centuries. Artificial Intelligence, as a digital humanities tool, enhances this historical inquiry by bridging archaeology, epigraphy, theology, and economic history. Whether through the recovery of lost inscriptions, mapping of trade routes, or computational analysis of religious texts, AI significantly deepens our ability to unearth and reinterpret the Buddhist past of Kashmir.

REFERENCES

- Agni Purana, translated by N. Gangadharan, New Delhi: Motilal Banarasidas Publication, 1954. [↵](#)
- Annual Report of the Archaeological Survey of India, Frontier Circle, for 1908-09.
- Archaeological Survey of India, Annual Report 1908-09, Frontier Circle, Calcutta: Superintendent of Government Printing, 1912. [↵](#)
- B.N. Puri, *India Under the Kushanas*, Bombay: Bharatiya Vidya Bhawan, 1965.
- Hiuen Tsang, *Si-Yu-Ki: Buddhist Records of the Western World*, Vol. 1, Translation from the Chinese of Hiuen Tsiang (CE 629-645) by Samuel Beal, London: Kegan Paul, Trench, Trubner & Co. Ltd, 1884.
- Kalhana, *Rajtarangini*, Vol. 1, translated from original Sanskrit by M.A. Stein, Westminster: Archibald Constable and Company, Ltd, 1900.
- Kalhana, *Rajtarangini*, Vol. 2, translated from original Sanskrit by M.A. Stein, Westminster: Archibald Constable and Company, Ltd, 1900.
- Lalita-Vistara, translated from the Original Sanskrit by Rajendralala Mitra, Calcutta: J.W. Thomas, Baptist Mission Press, 1881.
- M.M. Sylvain Lévi, and Édouard Chavannes. (1895). "L'itinéraire d'Ou-K'ong (751-790)." In: *Journal Asiatique*, (1895) Sept.-Oct., (In French), pp-358. [↵](#)
- M.A. Stein, Notes on Ou-K'ong's Account of Kashmir, Wien: Germany, 1896. [↵](#)
- Nilamat Purana, translated by Ram Lal Kanji, Lahore: Motilal Banarasidas Publication, 1924.
- Nilamat Purana: Translation by Ved Kumari from original Sanskrit, Delhi: Motilal Banarasidas Publication, 1968.
- R.K. Mookherjee, *Ashoka*, London: MacMillan and Co. Ltd, 1928. [↵](#)
- S. Radhakrishnan, *The Dhammapada*, Oxford: Oxford University Press, 1950. [↵](#)
- Ved Kumari, *Nilamat Purana: Translation from Original Sanskrit*, Delhi: Motilal Banarasidas Publication, 1968.
- Vishnu Purana, translated by H. H. Wilson, Chapters 17 and 18, publisher: Unknown, 1840. [↵](#)

Augmentation, Not Automation: Rethinking AI's Impact on Work in the Global South

Kuhu Singh

London School of Economics, London, UK

DOI: [10.1201/9781042018116-14](https://doi.org/10.1201/9781042018116-14)

ABSTRACT: This chapter examines the evolving relationship between artificial intelligence and the world of work in the Global South. Moving beyond automation narratives, it focuses on how AI can augment human capabilities and improve the quality of work, especially in contexts shaped by informality, resource constraints, and limited digital access. The chapter introduces a three-part augmentation framework that evaluates AI deployments by their effect on task structure (elimination, streamlining, reskilling), worker dignity (autonomy, recognition, stress), and alignment with Indian civilisational values such as Sahyog (cooperation), Seva (selfless service), and Sarvodaya (upliftment of all). Drawing on a comparative case study of AI tools in health, agriculture, and customer service, the chapter codes each deployment along these dimensions to identify what drives or constrains pro-worker outcomes. The findings show that dignity gains and promotion of civilisational values are possible but not guaranteed. The chapter concludes with policy and design recommendations to guide the development of inclusive, pro-worker AI systems in the Global South.

Keywords: Artificial Intelligence, Job Augmentation, Global South, Ethical AI, Pro-worker Technology, Labour Transformation

1 INTRODUCTION

In rural Rajasthan, community-health workers known as ASHAs now turn to a WhatsApp chatbot known as ASHABot, for instant guidance on everything from vaccine schedules to managing family disputes. The tool searches a doctor-curated knowledge base and employs Generative Pre-Trained Transformer 4 (GPT-4) to generate an answer from the retrieved information, to respond to questions sent by ASHAs ([Ramjee *et al.* 2024](#)). Since its pilot launch in 2024, 869 ASHAs have exchanged more than 24,000 messages with the bot, and participants reported faster response times and greater confidence when handling sensitive cases, registering tangible boosts to productivity and professional dignity ([Microsoft Research 2025](#)).

Across the Global North, discussions around the impact of artificial intelligence (AI) at work still open with stark forecasts about automation creating mass job displacements ([Frey and Osborne 2017](#); [The Future of Jobs Report 2025](#)). Yet ASHABot and other applications hint at a different story playing out in the Global South: augmentation rather than replacement. In contexts shaped by informal labour markets, limited digital infrastructure, and skills gaps, AI's most immediate and meaningful value lies in amplifying human capability rather than replacing it.

The existing scholarship in this area focuses on three different views. Automation-risk forecasts use task-exposure indexes to predict large-scale displacement. The figures vary, with Arntz, Gregory and Ziehran (2016) estimating nine per cent of jobs in OECD economies to be at risk of being automated, whereas other studies warn that up to 47 per cent of US employment is susceptible to computerisation ([Frey and Osborne 2017](#)). [Acemoglu *et al.* \(2022\)](#) find that companies exposed to AI are increasing hiring in AI-related roles while simultaneously reducing hiring in non-AI roles and changing the skill requirements of remaining positions.

A brighter picture is painted by studies of the impact of generative AI on productivity. [Brynjolfsson, Li and Raymond \(2023\)](#) find that the use of AI tools by customer support agents raised their productivity by 14% on average, with novice and low-skilled workers experiencing significantly higher productivity gains as compared to experienced and skilled workers.

[Noy and Zhang \(2023\)](#) find similar increases in productivity and reduced inequality, in the context of writing tasks. However, these studies focus almost exclusively on high-bandwidth corporate settings in the Global North.

The third strand offers a cautionary counterweight. Ethnographic studies of gig-drivers, warehouse pickers and online data-annotators show that AI systems can also act as “digital supervisors,” tightening surveillance and narrowing workers’ room for discretion ([Gray and Suri 2019](#)). These narratives remind us that efficiency gains may come at the cost of dignity, yet they rarely pin down how big the trade-off is or which design tweaks could soften it. Meanwhile, high-level ethics frameworks such as UNESCO’s [Recommendation on the Ethics of Artificial Intelligence \(2021\)](#) or NITI Aayog’s *Responsible AI* guidelines (Approach Document for India Part 1 – Principles for Responsible AI 2021), list ideals such as transparency and fairness but give scant advice on turning those ideals into everyday, measurable practice in the workplace.

No study yet combines bottom-up evidence from the Global South with a normative, values-based lens that highlights dignity, cooperation, service and upliftment of all. We still lack an answer to one central question: *How, and under what conditions, can generative and agentic-AI tools raise both productivity and worker dignity in developing-economy workplaces?* That is the gap this chapter aims to address.

We proceed as follows. Section 2 clarifies key terms (generative, agentic and predictive AI) and offers an “augmentation framework” that adds dignity and civilisational values to the usual efficiency metrics. Section 3 details the methodology; Section 4 presents the case descriptions; Section 5 presents the findings; Section 6 discusses implications and translates them into policy and design recommendations; and Section 7 concludes with a call for India-led, pro-worker AI standards.

By grounding the global debate in concrete frontline deployments and analysing them through a values-based lens, this chapter recasts AI not as a threat to labour, but as a catalyst for human-centred progress.

2 CONCEPTUAL FOUNDATIONS

AI systems affect occupations differently depending on whether they generate content, predict outcomes, or act autonomously; a clear typology therefore helps map which tasks are at risk, which are ripe for augmentation, and which require new safeguards.

Predictive or discriminative models (such as credit scoring, demand-forecasting, vision-based quality checks, etc) turn data into probabilities or classifications and have done so for years ([Jordan and Mitchell 2015](#)). *Generative* models build on that foundation to create new text, code, or images from a single prompt, increasing the speed of writing, design, and analysis tasks ([World Economic Forum 2024](#)). *Agentic* systems go one step further: they use a generative core plus planning, tool calls, and short-term memory to execute multi-step workflows (such as booking travel, resolving Tier-1 customer tickets, generating and testing software patches, etc) with minimal supervision ([Schneider 2025](#)). Because agentic autonomy depends on generative reasoning, these layers form a stack rather than three silos, and each tier changes labour in distinct ways. Prediction streamlines decisions, generation accelerates content creation, and agency can partially reallocate whole job fragments. Understanding this stack is essential for gauging which roles face displacement and which stand to gain from augmentation. This typology not only helps clarify what AI systems can do, but also lays the foundation for understanding how these systems reshape work in practice. To evaluate these shifts meaningfully, we turn to a three-part augmentation framework.

2.1 *Augmentation framework*

AI can impact work in three distinct ways. First, there are certain tasks which can be entirely performed by AI. These tasks are generally repetitive in nature or involve pattern detection. In such cases, AI would eliminate the human need to do those tasks, resulting in workers having more time. Second, AI can help speed up tasks that still require human involvement. For example, language models can help draft emails, and design tools can create quick visual mock-ups. The worker still guides the process, but things get done faster and often better. Third, AI can transform how entire workflows are carried out. Advanced systems can plan, use external tools,

and return detailed outputs such as a full software update or a ready-to-book travel plan. In these cases, the human role shifts toward supervision, decision-making, and creative thinking, which may require learning new skills but can also lead to higher-value work. This forms the first part of the augmentation framework, namely the way in which AI impacts the work flow itself.

While productivity metrics offer valuable insight into the efficiency gains enabled by AI, they do not capture the full impact on the nature of work. It is equally important to consider the quality of the worker's experience, particularly in terms of dignity and deeper civilisational values rooted in India's social and philosophical traditions. To assess this, each type of AI-driven change is evaluated across three dignity-related dimensions, forming the second part of the augmentation framework:

1. Autonomy – the degree of control the worker retains over task execution and interaction with the AI system;
2. Recognition – whether the organisation formally acknowledges the enhanced scope of the worker's contribution, through compensation, titles, or informal status;
3. Stress balance – the extent to which AI use reduces or intensifies cognitive load and feelings of surveillance.

In addition to dignity, long-standing Indian civilisational values offer a deeper normative lens for assessing the role of AI in the workplace. Three enduring Indian principles (*Sahyog*, *Seva*, and *Samatva*) are particularly relevant in this context. These form the third part of the augmentation framework.

1. *Sahyog* (cooperation): Embedded in traditions ranging from local panchayats to national cooperatives, *Sahyog* reflects a long-standing emphasis on collective decision-making. AI systems that align with this value actively involve workers in the design and deployment process by inviting feedback, enabling course correction, and ensuring that the benefits of efficiency gains are shared across the organisation.
2. *Seva* (selfless service): Work has historically been viewed not just as a means of livelihood, but as a contribution to the well-being of others.

AI tools that reflect the spirit of Seva prioritise use cases that reduce drudgery, support essential services, and extend benefits to those in underserved roles, reinforcing the idea that technological progress should serve broader social goals.

3. *Sarvodaya* (progress of all): Social and economic progress must benefit every member of society, especially the least advantaged. It frames well-being not as individual gain but as collective flourishing. A policy or technology is therefore judged by how far it advances the welfare of all, leaving no one behind.

Together, these values provide a normative foundation for evaluating augmentation. An AI intervention cannot be considered true augmentation if gains in productivity come at the cost of diminished dignity or a decline in core civilisational values. However, if productivity improves without negatively affecting these dimensions, the intervention may still qualify as augmentation. Ultimately, the goal is not just to enhance output, but to do so in a manner that upholds the integrity, autonomy, and well-being of the worker. By embedding these values into the design and deployment of AI, augmentation becomes not just a technical shift, but a pathway toward more equitable, human-centred progress.

3 METHODOLOGY

This study employs a document-based comparative case study approach, relying on publicly available sources such as policy documents, government reports, and secondary literature. Three Global-South deployments in healthcare (ASHABot), agriculture (Plantix), and customer services (GPT co-pilot in Philippine call centres) were selected because each (i) documents measurable AI use, (ii) represents a distinct sector, and (iii) offers publicly available evidence.

While it does not involve field interviews, the cases are selected for their relevance to emerging AI deployment patterns. Each case was coded deductively for task shift, dignity metrics and values alignment. Limitations of this methodology are that desk-based data means results are illustrative,

not causal; nonetheless they illuminate real augmentation pathways that large datasets miss.

3.1 *Coding frame*

Each case is coded along five pro-worker dimensions: task shift (eliminate/streamline/reskill), autonomy, recognition, stress balance, and alignment with the three Indian values (*Sahyog, Seva, Sarvodaya*).

4 CASE DESCRIPTIONS

Context- Accredited Social Health Activists (ASHAs) in rural India serve hundreds of families but have limited access to supervisors or up-to-date reference material. Routine queries regarding vaccination schedules, warning signs, family-planning advice often require time-consuming phone calls or paper manuals. Concerns about supervisors criticising perceived knowledge gaps often discourage ASHAs from seeking clarification, especially on sensitive issues such as sexual and reproductive health ([Ramjee et al. 2024](#)).

Intervention- ASHABot is a WhatsApp chatbot that combines a doctor-curated knowledge base with a GPT-4 language model. ASHAs send voice or text questions and receive instant, guideline-compliant answers, even on low-bandwidth connections. In its 2024 pilot, 869 ASHAs exchanged more than 24,000 messages; and users reported quicker decision-making and greater confidence when handling sensitive cases ([Microsoft Research 2025](#)).

4.1 *Plantix - AI crop-diagnosis app*

Context- Small farmers frequently lose yields to pests and diseases while waiting for an extension officer or agro-dealer to identify the problem with their crops. Rapid, accurate diagnosis is a persistent service gap.

Intervention- Plantix is a free smartphone application that applies image-recognition AI to more than 800 crop disorders ([Groupe Spécial Mobile Association 2024](#)). Farmers photograph an affected leaf and receive a diagnosis and treatment options within seconds; an in-app forum lets users compare remedies and share local tips. The app has surpassed 10 million downloads on Google Play and is now used across Asia, Africa, and Latin America. It can process up to 250,000 images daily with an accuracy of above 90% ([Groupe Spécial Mobile Association 2024](#)).

4.2 *GPT co-pilot – Customer support*

Context- Large call-centre teams handle high volumes of customer queries; novice agents typically lag behind experienced colleagues in speed and script accuracy, widening performance gaps and prolonging after-call work. Intervention- A real-time co-pilot built on a GPT-based language model listens to each call, suggests context-specific replies on the agent's screen, and auto-drafts case notes. In a 2023 staggered deployment to 5179 agents, the system raised resolved tickets per hour by 14 percent overall and by 34 percent for the least-experienced staff, while maintaining customer-satisfaction scores and reducing after-call documentation time ([Brynjolfsson et al. 2023](#)).

5 FINDINGS

5.1 *Cross-case patterns*

All three deployments begin by streamlining routine sub-tasks. Two of the deployments (ASHABot, GPT co-pilot) also trigger reskilling, as workers learn to curate or interpret AI output. Autonomy consistently rises when users initiate the AI interaction themselves and keep decision control. Stress falls where the tool removes sources of embarrassment (for ASHABot) or hostile customer exchanges (for customer service AI copilot). Recognition is uneven; only the call-centre links AI gains to formal incentives. Sahyog

and Seva appear in the two public-good sectors; Sarvodaya (equitable upliftment) is clearest in the BPO study, where novices benefit most, and also appears in the Plantix case study.

Taken together, the three cases show that gains in worker dignity are contingent, not guaranteed. Where frontline users help shape the tool, autonomy rises; where system logic is transparent and reduces cognitive or emotional burden, stress falls; and where organisations realign incentives, recognition follows. Equally important, each deployment demonstrates that AI can reinforce, rather than erode, core civilisational values such as co-operation (*Sahyog*), selfless service (*Seva*), and universal upliftment (*Sarvodaya*).

6 IMPLICATIONS AND DESIGN PRINCIPLES FOR PRO-WORKER AI

Together, the three case studies confirm that AI deployments can uplift worker dignity and reinforce civilisational values, but only under specific conditions. These gains are neither automatic nor universal; they emerge when systems are participatory, transparent, and anchored in shared benefits.

First, augmentation begins with streamlining but reaches its maximum potential through reskilling. The AI co-pilot case shows that when workers are empowered to curate and refine AI outputs, they transition from execution to supervision, resulting in higher learning velocity, task ownership, and recognition. Such transitions should be designed intentionally, not left to chance.

Second, autonomy is most reliably increased when AI interaction is user-initiated, not imposed. Both ASHABot and Plantix succeed because frontline workers control the point of use. Autonomy is not only a functional outcome but a psychological one linked to confidence, reduced fear of supervision, and greater trust in the tool. Design principles must therefore prioritise user control and opt-in pathways.

Third, gains in recognition require explicit incentive realignment. While AI improved productivity across all three deployments, only the call-centre

example translated those gains into formal organisational rewards. This highlights a critical policy gap: without structural mechanisms that link dignity to incentives, augmentation risks becoming invisible or extractive.

Fourth, stress balance improves most when AI systems reduce ambiguity or social friction. ASHAs benefitted from not having to “perform” knowledge in front of seniors; call-centre workers reported fewer high-conflict interactions with customers. This suggests that emotional and relational labour should be a core consideration when evaluating AI's impact on workplace well-being.

From a civilisational perspective, the values of *Sahyog* (cooperation), *Seva* (service), and *Sarvodaya* (upliftment of all) serve as normative filters for identifying what truly counts as pro-worker AI. AI that fosters *Sahyog* invites feedback and evolves through participatory design. AI that achieves *Sarvodaya* benefits the least advantaged through language access, interface simplicity, or disproportionate gains for novices. AI systems that embody the value of *Seva* are not defined solely by their technical design, but also by the context in which they are deployed. The ethical alignment of an AI tool depends on both how it functions and where it is situated. For instance, ASHABot reflects the spirit of *Seva* not merely because it reduces cognitive load or provides accurate information, but because it supports essential, community-oriented healthcare work in underserved rural areas. If the same system were repurposed to assist corporate interviewers in evaluating job candidates, the value alignment would shift, as the deployment context lacks the same service-oriented orientation. Thus, the sectoral application of AI plays a critical role in determining whether an intervention truly advances collective well-being or reinforces extractive hierarchies. Recognising this interplay between technical design and institutional setting is essential for ensuring that AI genuinely reflects and advances civilisational values.

To advance pro-worker AI, coordinated efforts are required across policy, institutional, educational, and technical spheres. Governments should link AI development grants and subsidies to participatory design requirements and develop sector-specific dignity scorecards to ensure transparency and public accountability. Employers can institutionalise worker voice by establishing internal AI councils and revising incentive structures to reward not only productivity but also gains in autonomy, skill acquisition, and

stress reduction. Educational and skilling institutions should introduce micro-credentials tailored to AI-supervised workflows and support peer-learning communities that help workers adapt collectively. Meanwhile, AI developers and vendors must prioritise low-bandwidth, multilingual interfaces to ensure inclusivity, and expose model reasoning through decision logs to build trust and enable user oversight. Taken together, these measures affirm that pro-worker AI is not merely a technological upgrade but a socio-technical ecosystem whose success depends on participatory design, value alignment, and equitable distribution of outcomes. Dignity, in this vision, is not an optional add-on but a foundational design constraint.

7 CONCLUSION AND LIMITATIONS

This chapter has sought to reframe the global AI-at-work debate through the lens of augmentation rather than automation, particularly in the context of the Global South. Drawing on examples from India and beyond, we show that AI can enhance worker autonomy, reduce stress, and affirm civilisational values, provided it is deployed with intention and equity.

The proposed augmentation framework, incorporating dignity metrics and normative filters such as Sahyog, Seva, and Sarvodaya, enables a more holistic evaluation of AI's workplace impact. It invites policymakers, developers, and employers to move beyond output metrics and consider the worker's full experience – technical, emotional, and ethical.

Yet, several limitations remain. First, this study is based on desk research and publicly available data; in-depth fieldwork would provide richer insights into workers' lived experiences. Second, our focus on successful deployments may understate the risks of opaque or top-down AI systems. Third, while the Indian civilisational lens offers unique normative strength, future work should explore resonances and tensions with other cultural value systems.

Despite these limitations, the chapter affirms a key possibility: that the Global South can lead in designing inclusive, dignity-centred AI futures. As India positions itself as both a digital powerhouse and a philosophical

torchbearer, it must ensure that the AI revolution it helps build is not only efficient but just and equitable.

REFERENCES

- Acemoglu, D, Autor, D, Hazell, J, and Restrepo, P (2022), ‘Artificial intelligence and jobs: Evidence from online vacancies’, *Journal of Labor Economics*, vol. 40, no. S1, pp. S293–S340, DOI:[10.1086/718327](https://doi.org/10.1086/718327).[↵]
- Approach Document for India Part 1 – Principles for Responsible AI (2021), NITI Aayog, viewed 26 July 2025 (<https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf>).
- Arntz, M, Gregory, T, and Zierahn, U (2016) ‘The risk of automation for Jobs in OECD countries: A Comparative Analysis’, *OECD Social, Employment and Migration Working Papers*, No. 189, DOI:[10.1787/5jlz9h56dvq7-en](https://doi.org/10.1787/5jlz9h56dvq7-en)
- Brynjolfsson, E, Li, D, and Raymond, L R (2023), ‘Generative AI at Work’, NBER Working Paper, no. 31161, viewed on 27 July 2025, (<https://www.nber.org/papers/w31161>).[↵]
- Frey, CB and Osborne, MA (2017), ‘The future of employment: How Susceptible are jobs to computerisation?’ *Technological Forecasting and Social Change*, vol. 114, pp. 254–280.[↵]
- Gray, ML, and Suri, S (2019), *GHOST WORK: How to Stop Silicon Valley from Building a New Global Underclass*, Houghton Mifflin Harcourt, viewed 25 July 2025, (<http://sarkoups.free.fr/gray2019.pdf>).[↵]
- Groupe Spécial Mobile Association (2024). *Detecting and Managing Crop Pests and Diseases with AI: Insights from Plantix*, Groupe Spécial Mobile Association, viewed 28 July 2025, (<https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/blog/detecting-and-managing-crop-pests-and-diseases-with-ai-insights-from-plantix/>).[↵]
- Jordan, MI, and Mitchell, TM (2015), ‘Machine learning: Trends, perspectives, and prospects’ *Science*, vol. 349, pp. 255–260, viewed 25

July 2025, (<https://www.cs.cmu.edu/~tom/pubs/Science-ML-2015.pdf>).[↗]

Microsoft Research (2025), *How ASHABot empowers rural India's frontline health workers*, Microsoft Research, viewed 26 July 2025, (<https://www.microsoft.com/en-us/research/story/how-ashabot-empowers-rural-indias-frontline-health-workers/>).[↗]

Noy, S and Zhang, W (2023), 'Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence' *Science*, vol. 381, pp. 187–192, DOI:[10.1126/science.adh2586](https://doi.org/10.1126/science.adh2586)[↗]

Ramjee, P, Chhokar, M, Sachdeva, B, Meena, M, Abdullah, H, Vashistha, A, Nagar, R, and Jain, M (2024), 'ASHABot: An LLM-Powered chatbot to support the informational needs of community health workers', *CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, article no. 698, pp. 1–22, <https://doi.org/10.48550/arXiv.2409.10913>[↗]

Recommendation on the Ethics of Artificial Intelligence (2021), *United Nations Educational, Scientific and Cultural Organisation*, viewed 25 July 2025, (<https://unesdoc.unesco.org/ark:/48223/pf0000381137>).[↗]

Schneider, J (2025), 'Generative to Agentic AI: Survey, Conceptualization, and Challenges' arXiv preprint, viewed 26 July 2025, <https://doi.org/10.48550/arXiv.2504.18875>[↗]

The Future of Jobs Report (2025), *World Economic Forum*, viewed 26 July 2025, (https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf).[↗]

World Economic Forum (2024), 'Jobs of Tomorrow: Large Language Models and Jobs', White Paper, viewed 27 July 2025, https://www3.weforum.org/docs/WEF_Jobs_of_Tomorrow_Generative_AI_2023.pdf.[↗]

From Tols to AI in Indian Education: Addressing Educational Inequality

Saurabh Kumar Shanu and Anu Singh

University of Delhi

DOI: [10.1201/9781042018116-15](https://doi.org/10.1201/9781042018116-15)

ABSTRACT: The chapter traces the historical roots of exclusivism in the Indian education system, from pre-colonial to contemporary times. It examines how traditional institutions like Tols, reserved for Brahmins, and Pathshalas, open to broader groups, were shaped by caste-based segregation. The colonial introduction of Western education, rather than disrupting these hierarchies, further entrenched elitism by privileging the urban, upper-caste elite. Post-independence, the expansion of public and private institutions brought new opportunities, but disparities in quality persisted, driven largely by economic and social inequalities. The poor and marginalized remained excluded from access to quality education. The chapter concludes by exploring the rise of Artificial Intelligence (AI) in education. While AI holds the promise of democratizing learning and expanding access, it also risks deepening existing divides if its deployment remains limited to privileged groups. By situating AI within India's historical and structural context, the chapter raises a critical question: will AI disrupt educational exclusivism, or reinforce it in new, technological forms?

Keywords: Indigenous Education, Colonial Education, AI in Education, Digital Divide, Educational Inequality

1 METHODOLOGY

This study adopts a qualitative historical–structural methodology combining archival analysis with critical policy review. It traces continuities and ruptures in educational access from indigenous institutions to contemporary AI-driven models, using secondary literature and official reports. A critical interpretive framework is applied to assess whether AI functions as a tool of educational democratization or as a new modality of exclusion.

2 INTRODUCTION

Artificial intelligence (AI) through personalised learning pathways, real-time assessment system, and scalable pedagogical innovations is fundamentally restructuring the global educational landscape. India which faces entrenched structural flaws — such as under-resourced schools, teacher shortage, rigid curriculum — AI holds immense transformative potential that can translate into meaningful educational outcomes to address the constraints. It becomes imperative to understand the historical evolution of India's entrenched structural flaws in the education sector which despite with all its achievements and contributions has a legacy of being exclusive.

The evolution of education in India is a reflection of a complex heritage of institutional continuity and structural transformation. A unique feature of this legacy is the persistence of exclusionary patterns across the historical epoch. While there have been adaptations of new technologies and ideological frameworks educational exclusivism through systematic social and economic privileges of certain groups has been a consistent feature of the educational landscape of India. The pre-colonial system of education in India was characterised by institutionally embedded caste hierarchies where some institutions served primarily Brahmins students with Sanskrit as the medium of instruction ([Naik and Nurullah, 1974](#)). While some of the scholars like Dharampal paints a romantic picture of pre-colonial India where the ‘content of the studies was better than what was then studied in England’ and the only flaw being its inability to provide the similar opportunities to women in terms of education (2000, p. 20) others like

Cheney, Ruzzi and Muralidharan have contested these claims by calling the traditional Hindu education ‘tailored’ for the requirements of Brahmin students by the Brahmin teachers (2005, p.1).

The solidification of the colonial regime marked a decisive transformation in the domain of education in India. The advent of Western education in India was profoundly shaped by the works of Charles Grant and T.B. Macaulay who emphasised on the introduction of English-medium instruction over the vernacular learning with explicit aim to create a ‘a class of persons, Indian in blood and colour, but English in taste, opinions, morals and intellect’ ([Macaulay, 1835](#), p. 115). The growth of western education institutions such as Presidency College, Calcutta Sanskrit College and numerous missionary schooling institutions established novel institutional hierarchies that instead of replacing overlaid the traditional caste-based exclusivity in education. The colleges like Hindu College of Calcutta had limited social inclusivity as it catered predominantly to the upper-caste Hindu Bhadrakalok youth consolidated existing caste and class hierarchies ([Suman and Shanu, 2024](#)). The elitist and exclusionary characteristics embedded within the colonial educational vision has also been emphasised by Krishna Kumar who argues that the colonial education was mostly picked by the upper castes which further enabled them to morally and intellectually place above the illiterate masses of India (2014).

Colonial education in India never truly functioned as a neutral engine of ‘modernity’ as it seldom led to a fracture from the traditional hierarchies. Sabyasachi Bhattacharya's edited Volume *Education and the Disprivileged* affirmed that the school system designed under British rule largely reproduced — rather than dissolved — older hierarchies of caste, class and gender. Separate hostels, waterpots, and examination rooms for the ‘Untouchables’ were provided to cater the sentiments of the Hindu Upper caste further reinforcing the notion of pollution at the institutions of education. Girls’ schools that did open taught sewing, hygiene and Bible stories. It was an irony that the places to liberate human mind and thought became a centre of reinterlocking caste and gender divide (2002). Feminist and subaltern works of Rosalind O’Hanlon and Padma Anagol have revealed that schooling for girls, when it existed at all, were rigidly supervised and aimed at training women for domestic chores. O’Hanlon in her critical work on Tarabai Shinde's *Stree Purush Tulana* affirms the

agenda of colonial jurisprudence for keeping women uneducated and economically dependent. Dalit girls rarely crossed the school threshold. Even where late-nineteenth-century missionary institutions admitted them, they were channelled into Bible-women or midwifery courses with no clear wage ladder (1994). Padma Anagol on the other hand has explored the concept of 'Navin Stri' (New Women) in India of who pursued secondary and university education in Bombay and Pune during 1850-1920. Her research highlights the Brahmin and upper-middle-class dominance at the English girls' high school and the explicitly domesticated curricula emphasising childcare, household hygiene and reinforced gendered division of intellectual labour (2006).

The colonial India also witnessed transformative subaltern counter-movements by intellectual and reforms like Jyotirao Phule, Savitribai Phule and Dr. B.R. Ambedkar who attempted to challenge the deeply entrenched exclusiveness of colonial and Brahmanical education structures. However, the Phule couples and Ambedkar faced both organised resistance and institutionalised discrimination from the society. Despite limited resources and brutal opposition their efforts became foundational to post-independent India's constitutional commitment for educational equality.

The Indian Constitution lays down a robust mechanism for equality of education in India. However, the qualitative have continued despite the constitutional measures as the quality gap in private and public school has intensified. The constitution establishes a comprehensive legal framework for educational equality including Article 21A (Fundamental Right to Free and Compulsory Education for all Children between the age of 6 and 14); Article 14 (Right to equality); Article 29.2 (prohibiting discrimination in admission to educational institutions maintained by the state on grounds of religion, race, caste or language).

In the context of the above, this chapter argues that Indian education has consistently reproduced structures of exclusion through changing institutional arrangements. Within this historical continuum, AI is not treated as a radical rupture but as a contingent technological moment whose capacity to advance educational equality is conditional upon robust, state-led, and rights-based regulatory and redistributive interventions.

3 INDIGENOUS INSTITUTIONS OF LEARNING: PRE-COLONIAL REALITIES

The traditional knowledge in India has been transmitted orally before being preserved in written form. Indigenous learning, which concentrated on the study of Indigenous history, culture, and values, was a crucial part of this larger knowledge system ([Kincheloe, 2006](#)). The objective of indigenous learning is to increase awareness of and appreciation for the experiences that Indigenous peoples have experienced. It is made possible by traditional institutions of learning such as Madrasas, Pathshalas, and Tols ([Hossain, 2024](#)). The educational landscape of pre-colonial 19th century Indian has been broadly classified into two distinct and often interrelated educational traditions which catered to the religious, cultural and vocational requirements of Hindu and Muslim communities. Pathshalas and Tols were the key institutions of Hindu education system and Madrasas and Maktabas were the outcome of Islamic traditions in India.

Pathshalas were a kind of elementary schools serving to a broad range of students. They provided a foundational space for early learning within Hindu society. The medium of education was used to be the vernacular language of the region ranging from Bengali, Magahi, Hindi in northern India ([Adam, 1868](#)). Pathshalas generally provided instruction in commercial and agricultural accounting practices with students practising their lessons on naturally available mediums such as on ground using a stick or finger, wooden boards, or slates. In addition to basic education, they also provided moral teachings to the children through texts like *Dan Lila* and *Dadhi Lila* which described the youthful amusements of Lord Krishna. Other significant moral texts included the works like *Sudama Charita*, *Ramcharitramanas* and *Sundar Kand*. These all texts were inspired from the Ramayana. The unique feature of Pathshalas was that it catered to a diverse range of students coming from different caste groups including students from marginalised and socially stigmatised castes like *Dosads*, *Pashis* and *Luniars*. These institutions had evolved in response to the localised requirements and more specifically to the educational requirements of a small number of families shaped by expertise and

knowledge of individual *gurus* (teachers) to address the unique need of their students ([Buchanan, 1812](#), pp. 297-302).

Tols were, however, a specialised category of Hindu educational institutions focused on the advanced study of traditional sciences and Sanskrit literature. The Pathshalas and Tols were not interlinked and interdependent as the latter operating independently only admitted students coming from the Brahmin community. The curriculum of Tols was also broader and exclusive providing rigorous education in Sanskrit, Religious texts, philosophical doctrines that were central to Hindu intellectual and spiritual life ([Naik and Nurullah, 1974](#), pp. 24-25). The Tols were ran by Brahmins and the syllabus included disciplines like Vedas, the scriptures on the major six schools of Hindu philosophy (Samkhya, Yoga, Nyaya, Vaisheshika, Purva Mimamsa and Vedanta), astrology, literature, lexicology, and logic. The *guru* was viewed as an ultimate authority with authority over the discipline. In the 18th century, it was normal to respect the *guru* and his wisdom, while also expecting disciples to be faithful to him ([Nawani and Jain, 2010](#)).

On the contrary, both the institutions of Muslim education were interconnected and inter-related. Maktabas were the institutions of primary education, a type of elementary schools, typically linked and affiliated with a mosque or a Muslim shrine. The primary objective of Maktabas was to provide foundational teachings of Islam through Quranic memorisation, however, in India the institution of Maktab evolved to accommodate specific local and socio-economic realities. Unlike the Maktabas of other Islamic regions where Arabic used to be medium of instruction Indian Maktabas adopted Persian to be the language of teaching due to Persian being the language of administration, culture, and Islamic scholarship in Mughal India and the later successor states ([Khan, 1983](#)). Unique to the Indian subcontinent, the teacher of the Maktabas was not essentially required to be a Muslim and Hindu teachers also taught Persian. *Kayasthas* (an influential Hindu social group primarily engaged in administration, writing and accounting) were the predominant group who studied as well as taught at the Madrasa allowing them to secure positions within the Mughal bureaucracy and earning them the title of *Monashee* (accountant) ([Bellenoit, 2012](#), p. 352). For instance, there were 865 Hindu students in the Maktabas

of Patna region and among them 711 belonged to the *Kayastha* caste ([Adam, 1868](#), p. 205).

Madrasas were the institution for secondary and higher education within the Muslim institutions of learning for the advancement of scholarship. The institution was built upon the foundational knowledge provided by Maktabas ([Alatas, 2021](#), p. 195). The Qur'an, the hadiths (sayings of Prophets), Fiqh (jurisprudence), Arabic, grammar, rhetoric, and logic usually constituted the curriculum. Madrasas often provided a comprehensive residential setting for the students where they received, along with advanced instruction, free lodging and meals. In certain cases, students were provided financial support through stipends in order to facilitate their academic and intellectual pursuit ([Khan, 1983](#)).

The pre-colonial educational landscape was characterised by inclusivity as well as marked exclusiveness where the specialised and higher education were open only to the socially dominant and influential communities such as Brahmins and Kayasthas. Although institutions like Pathshalas were open for all the communities, the curriculum was limited to providing foundational training in vernacular languages and more rigorous and advanced education with potential to provide social mobility remained an exclusive domain of a few.

4 EDUCATING THE FEW: COLONIAL EXPERIMENT WITH EDUCATION

The advent of western education in India under the colonial regime marked a departure from the indigenous system of education that had evolved over centuries under the socio-political realities of the region. The Colonial educational policies redefined the nature, content, and purpose of knowledge. They legitimised Western epistemologies and systematically marginalised indigenous traditions and local knowledge systems. Utilitarian and bureaucratic objectives to advance the narrow conception of 'institutional learning spaces', called for the restructuring of curriculum, which also aligned with the administrative requirements of the colonial state to facilitate their rule in India through English educated Indians. The

shift also entrenched and institutionalised rote learning and reconfigured pedagogical norms and privileged specific modes of instruction through Western schools and Colleges. The establishment of formal educational institutions, standardised textbooks, and teacher-centred classrooms fundamentally altered the roles of educators, learners, and teaching materials. Moreover, the colonial administration centralised educational governance within a few bureaucratic agencies, thereby shaping both the production and dissemination of knowledge. Despite frequent proclamations about the need to educate the colonised population, the colonial state allocated insufficient funding towards this end. There was a systematic neglect of the elementary schooling for the masses in favour of higher education for the elites coming from upper-caste and urban backgrounds ([Nawani and Jain, 2010](#), p. 505).

Rao has showed that, through Orientalist initiatives, the early colonial educational institutions such as Sanskrit College of Banaras (1791) which was founded by Jonathan Duncan under Lord Cornwallis's patronage, Sanskrit College of Calcutta (1821) established by Lord Amherst's administration, Hindoo College of Poona (1821) set up with the help of Peshwa's Dakshina Fund, and College of Fort St. George of Madras (1812) established in 1812 as a training centre for East India Company (EIC) servants exclusively employed or enrolled Brahmins only (2019, p. 90).

Charles Grant's Observations marked a notable shift from orientalist perspectives in the sphere of education in India. Motivated by racial theory Grant proposes that Indians have the potential for moral improvement through the dissemination of 'light and knowledge (1792, 148).' Grant utilised Orientalist scholarship to argue that despotism had always been inherent in Hindu nature, emphasizing their supposed weakness and indolence compared to the British. Grant's 'Observations' was also instrumental in the formation of the 1813 Charter Act for allocation of funds for Indian education and opening British territories to missionaries. The document was reprinted by the English Parliament in 1813 and provided the intellectual foundation for the evangelical campaign in education to transform Indian society ([Zastoupil and Moir, 2013](#), p.6). The racial theoretical framework established by Grant was later further developed by Thomas Babington Macaulay, whose 1835 Minute on Education employed similar arguments on Indian intellectual and moral

inferiority ([Bhabha, 1984](#)). Colonial interventions began with Macaulay's Minute of 1835, and Wood's Despatch of 1854 further replaced the vernacular indigenous learning with English-medium education. The educational philosophy of Macaulay inherently aimed to create 'a class of persons, Indian in blood and colour, but English in taste, in opinions, in morals and in intellect' (Macaulay, p.115). This renowned policy embodied the colonial strategy of producing a carefully selected intermediary class of Indians. The carefully constructed class was supposed to serve the British administrative needs and at the same time remain culturally cut off from the native masses who were also 'unaware and ignorant of the Indian culture and writing' ([Jana and Sarkar, 2021](#), p.6). English education became an essential tool for social mobility in colonial India, but its benefits were limited to the upper-caste groups and urban elites. For instance, in Bengal, the Bhadrakalok class—typically comprised of upper-caste Hindus—capitalised on the early establishment of Western-style institutions to secure dominant positions in administration, culture, and society. Access to these institutions was tightly linked to family wealth, urban residence, and social networks, which entrenched privilege intergenerationally ([Maitra, 2023](#)). Simultaneously, the Wood's Despatch of 1854 institutionalised a bifurcated educational framework that maintained English as the official medium of instruction at the university level along with promoting vernacular languages for primary education. It effectively entrenched a dual educational system segregated along social and linguistic lines. English-medium institutions were predominantly accessible to the socio-economic elites to serve as gateways for upward social mobility and administrative employment. In contrast, vernacular-medium schools, which predominantly catered to lower social strata, were systematically under-resourced and marginalised within the colonial educational apparatus. Consequently, this dichotomous structure further strengthened linguistic hierarchies. In the hierarchy, the English educated elites occupied a dominant position due to their cultural and social superiority. As a result, the English education further perpetuated educational inequalities and exacerbated social stratification ([Sajid and Hussain, 2025](#)).

Further, as the Government undertook the duty to mass educating people at the school levels, including the Untouchables, the upper castes fiercely opposed their entry into schools. The [Hunter Commission of 1882](#) notes

that the number of people coming from the lower strata was growing in the institutions run by the Government. The Report of the commission mentions in context of the province of Bombay, ‘In Bombay the greatest attention has been paid to the matter; whereas in 1871 there were only 592 low-caste boys at school, there are now 3,512 of whom 2862 are in departmental day school’ (Hunter Commission, 1882, p.149). But the Commission also acknowledges the discrimination they had to face at the places of education. The Commission further notes that in Bengal ‘the low caste, who sits on a separate mat, does not defile his neighbour and mat freely attend school ... the higher castes generally are strongly averse to their children mixing with low caste boys ... Madras witness mentions the case of a school for Cherumars, the ancient slave caste, being established at Calicut but the Nayars and Tiyars “used to waylay the boys as they went to school and snatch their books out of their hands.” A Bombay witness relates how some promising low-caste boys were recently sent from the regimental school at Dharwar to the Government high school, when a large number of Brahman boys seceded at once from the high school’ (Hunter Commission, p.514).

The colonial experiment, along with the creation of an English-educated elite, witnessed the permeation of caste-based discrimination into the sphere of education in a new form, giving rise to fresh practices of exclusion. This differed from pre-colonial discrimination in educational spaces, where the so-called low castes were simply denied access altogether at institutions like Tols.

5 POST-COLONIAL INDIA AND THE EDUCATIONAL ENDEAVOURS

India's first national assessment of literacy, conducted through the 1951 Census, recorded an overall literacy rate of only 18.33%, with a pronounced gender gap—27.16% among males and merely 8.86% among females ([PIB, 2003](#)). In response, the Indian Government undertook an unprecedented expansion of educational institutions, driven by constitutional commitments to universal education, equality, and social justice.

To begin with, the evolution of education in India can be broadly classified in four periods or phases viz., Early Foundation Period (1947-1964), Policy Institutionalisation Phase (1968-1986), Rights-Based Approach (1990s-2010s) and the Contemporary Reforms (2020-Present). The Early Foundation Period began with the University Education Commission (1948-49) led by Dr. S. Radhakrishnan which laid the foundation for restructuring the higher education in India. The Commission proposed that the objective of the higher education in the independent Indian should be to create a group of intellectuals 'who would bring together the two, material resources and human energies' ([Radhakrishnan Commission, 1949](#), p.18). The focus was to create universities as a hub where new ideas would be generated and those which are likely to impede the development of the country be discarded for the growth of a democratic 'New India' which believed in 'the inherent worth of the individual in the dignity and value of human life' (Radhakrishnan Commission, pp.18-20). The Radhakrishnan Commission was followed by the Secondary Education Commission of 1952-53 —Mudaliar Commission— to assess and improve the state of secondary education in the country. The Commission acknowledged the isolation of schools from life and limitations of the 'curriculum formulated and presented through the traditional methods of teachings' ([Mudaliar Commission, 1953](#), p. 21). The commission argued that the aim of secondary education should be to improve vocational efficiency, development of personality to create leadership for the future (Mudaliar Commission, pp. 22-29). It recommended mother tongue as the medium of instruction along with three language formula in the schools. However, there were also some elements of gender bias as the commission also prescribed a mandatory Home science course exclusively for women. Other subjects were common for both girls and boys. However, the two decades following the independence made little progress in education noting that 'the present system of education, designed to meet the needs of an imperial administration ... will need radical changes if it is to meet the purpose of a modernising democratic and socialist changes' ([Kothari Commission, 1966](#), p.5).

Kothari Education Commission of 1964-66 was the first to produce the most comprehensive report on education and provided foundation for India's national education framework. The commission while

acknowledging that ‘it is education that determines the level of prosperity, welfare and security of the people’ (1966, p.1) recommended a common schooling system of schools with ten years of primary followed by two or three years of lower secondary stages providing a general education (1966). Following the Kothari Commission, the National Policy on Education marked India's first systematic effort to build an integrated national system, advocating free and compulsory education up to age fourteen and a common school system to equalise opportunities across caste, class, gender, and religion (GOI, 1968). Further, the National Policy on Education (NPE) 1986 was launched with an emphasis that ‘In our national perception education is essentially for all’ (NPE, 1986, p.3). It envisioned a National System of Education with a common structure of 10+2+3 to be based on a common national curricular framework (1986, p.4). The Policy had an independent chapter ‘Education for Equality’ aspiring to remove ‘disparities and equalise educational opportunity by attending to the specific needs of those who have been denied equality so far’ (1986, p.6).

To operationalise the objectives of the 1986 National Policy on Education, the Programme of Action (1992) outlined concrete implementation strategies through a decentralised framework, including Operation Blackboard, while emphasising targeted interventions for women, Scheduled Castes, Scheduled Tribes, and educationally backward minorities ([GOI, 1992](#)). During the post-1990s economic reforms, neoliberal policy orientations expanded, with international agencies such as the World Bank influencing educational financing and reform agendas. Initiatives like the District Primary Education Programme (1994) targeted educationally backward regions to improve access and quality. Complementing these reforms, the Eighty-Sixth Constitutional Amendment Act (2002) constitutionally mandated free and compulsory education for children aged six to fourteen, rendering the Right to Education a justiciable fundamental right ([GOI, 2002](#)). The National Education Policy, 2020 is the most recent comprehensive reform which focuses on holistic and multidisciplinary education with integration of Indian Knowledge System and technology-enabled learning and flexible curricular structures where multiple entry and exit opportunities are provided to the students ([GOI, 2020](#)).

Table 1. Literacy Rate in India.

Period	Total Literacy	Male Literacy	Female Literacy	Gender Gap
1951	18.33%	27.16%	8.86%	18.30%
1981	43-75%	56.38%	29.75%	26.62%
2022	74.04%	82.16%	65.46%	16.68%

The census of 2011 recorded a four-fold increase in total literacy from 1951 with reduced gender gap. However, these aggregate figures mask the existing disparities and exclusions resulting out of limitations of post-independence policies. The structural inequalities in education persists across multiple dimensions viz., socio-economic, accessibility, infrastructural adequacy, geographical distribution, and identity-based discrimination. The liberalisation of economy in 1991 resulted in exponential growth of private education institutions in India where they proliferated to meet the market demand ([Velaskar, 2010](#)). The privatisation of education led to differentiated quality of learning depending on the nature of institutions viz., public or private. The access to quality education at the private institutions also became almost impossible for economically disadvantaged students with colleges/schools charging unaffordable fees. It created a parallel educational sphere where private institutions providing quality education served the affluent and the under-resourced public institution becoming the only viable options for the others. The private institutions also offered superior infrastructure, faculty and learning opportunities ([Muralidharan and Kremer, 2008](#)).

The infrastructure deficiencies in Indian schools are stark which make education inaccessible to students with special needs. The gaps encompass inadequate physical facilities, insufficient water, sanitation and limited technological integration with pronounced gaps in government and private schools. According to the Unified District Information System of Education (UDISE+) report, in 2023-24 only 33.3% of government schools had disabled-friendly toilets. As far as the computer facilities are concerned,

only 43.5% government schools had access to computer for pedagogical purpose compared to 70.9% in private schools ([GOI, 2024](#)).

Educational inequality is especially visible in the rural–urban divide. With around half of students finding simple mathematical division difficult, children in rural parts struggle with foundational learning skills in schools with poor infrastructure and funds. ([ASER, 2024](#)). The foundational learning deficit is further compounded by pronounced gender disparities where boys outperform girls in arithmetic and reading. Further, approximately one-quarter of adolescents lacked basic literacy, while unequal smartphone access which was significantly lower among females further constrained girls’ educational and digital competencies. The performance of male counterparts was also better than females across all task concerning digital awareness (2024).

Post-independence policy interventions have been able to improve the literacy rate, and access to schools, however, the exclusiveness of education continued, and relevance of educational provisions remained compromised in terms of quality, infrastructure, foundational skills and more especially digital infrastructure and literacy in the 21st century. The demands of the century require not merely expanded access but transformed educational experiences that develop critical thinking, digital literacy, and foundational skills across all demographic groups.

6 ARTIFICIAL INTELLIGENCE: TECHNOLOGICAL PROMISES AND INSTITUTIONAL ADOPTION

The transformative potential of Artificial Intelligence (AI) in education, primarily, rests on the technology's capacity to personalise learning experiences and provide quality education interventions across a diverse geographical and socio-economic landscape. AI can be utilised to democratise education by effectively addressing the persisting challenges pertaining to education system in India. They include student-teacher ration gap, limitations related to standardisation of curriculum along with differential learning needs (Dundhar *et al.*, 2014; [Misra and Ghadai, 2015](#)). Intelligent Tutoring Systems (ITS) is one of the computer-based educational

tools that uses AI to provide personalise and adaptive instruction to the students with providing immediate feedback to the students in order to promptly help them understand their mistakes and rectify them ([Lin, Huang, and Lu, 2023](#)). Other AI based tools such as, adaptive learning platforms, and automated assessment tools promise to democratise access to high-quality educational experiences regardless of geographic boundaries ([WEF, 2024](#)). The New Education Policy (NEP) 2020 explicitly acknowledges the significance of AI encourages for its integration for student development as it ‘will not just change what students learn in the classroom but how they learn’ (NEP, 2020, para. 23.2).

Artificial Intelligence (AI) has potential to transform education in India. It offers new pathways to democratise access to education, personalise learning, and enhance the efficiency of teaching–learning processes. Evidence from AI-driven multi-sensory literacy programme implemented in over 5,000 government schools across seven states has demonstrated tangible learning gains of 20–40% in English reading and comprehension ([Srinivasan and Murthy, 2021](#), p.1). The schools were resource constrained and despite having infrastructural limitations the AI based programmed showed positive results in foundational literacy. Beyond literacy, AI applications in India, as documented by Jaiswal and Arun, include personalised learning systems, adaptive assessments, and recommendation engines that tailor pace, content, and feedback to individual learners. It allows and enables remediation and enrichment (2021). AI can early identify at-risk students through predictive analytics, and it provide space for targeted interventions to reduce dropout risk and strengthen equitable educational growth ([Lopez, et.al, 2024](#)). This aligns with global policy visions, though Schiff’s cross-national analysis of AI strategies warns that in most countries, including India, policy discourse still prioritises “education for AI” (workforce training) over “AI for education,” and fails to adequately address the ethical and equity implications of AI in classrooms ([Schiff, 2021](#)).

The democratising potential of AI in India rests on its ability to transcend geographic and social barriers. Adaptive AI platforms, capable of functioning in multilingual contexts and accommodating diverse learning styles, could bridge the quality gap between urban elite and rural under-resourced schools ([Shalini and Tewari, 2020](#)). For students in remote areas,

AI-powered offline learning tools can mitigate the lack of trained teachers and standardised pedagogy, while automated grading, virtual tutoring, and real-time feedback free teachers to focus on higher-order skills and socio-emotional development ([Roy, 2020](#)). Moreover, AI's role in generating data-driven insights can support policy planning, as seen in Andhra Pradesh's machine-learning initiative to predict potential dropouts based on socio-academic indicators (Roy). These capabilities hold promise for scaling quality education, aligning with the Sustainable Development Goals by fostering inclusive, lifelong learning opportunities.

However, the transformative promise of AI is contingent upon resolving deep-seated structural and ethical challenges. India's school system, particularly its government sector, is marked by a severe digital divide—amplified during the COVID-19 pandemic when ASER's 2020 survey revealed that in rural India only about two-thirds of households had smartphones, with poorer and government school households disproportionately excluded ([ASER, 2021](#)). Limited internet connectivity, insufficient devices, unreliable electricity, and low digital literacy among teachers constrain AI adoption, particularly in rural and marginalised regions (Lopez, *et al.*). This raises the concern that AI might exacerbate, rather than reduce, inequities—privileging private schools and affluent students while leaving behind those in infrastructure-deficient government schools. The risk is not hypothetical; Shalini Tewari's work highlights that without targeted interventions, AI could reproduce existing hierarchies of access and quality, entrenching the exclusivity of elite education (Shalini and Tewari).

Ethical and governance gaps compound these infrastructural disparities. AI is a data-intensive technology, and in the absence of robust data protection laws with provisions for algorithmic accountability—India's current framework lacks explicit rights to opt out or be informed of automated decision-making—there are risks of privacy violations, opaque profiling, and misuse of sensitive student data (GISW, 2019). The governance vacuum in regulating AI in education means there is little oversight over algorithmic bias, transparency, or fairness, raising questions about whether AI tools are culturally responsive or inclusive of marginalised learners' needs. Schiff's policy review cautions that while national AI strategies globally discuss ethics in sectors like healthcare and

transport, there is a near-complete absence of sector-specific ethical guidelines for AI in education (2021). This omission is critical given that education shapes future opportunity structures and social mobility.

Even where AI tools are deployed, sustainability and contextualisation remain pressing challenges. A historically informed comparison reveals strong continuities between AI-enabled elite schooling and colonial English education in the nineteenth century. Just as English-medium institutions such as Hindu College and Presidency College functioned as gateways to administrative authority and social mobility for a narrow, urban, upper-caste elite, AI-integrated private and international schools will operate as sites where technological capital is converted into cultural and economic advantage. In contrast, government schools, much like vernacular institutions under colonial rule, may remain largely excluded from advanced technological interventions, reinforcing a structurally dual educational system. The present technological stratification also closely mirrors the colonial English–vernacular bifurcation, wherein access to power was mediated through language; today, it is mediated through digital infrastructure, algorithmic access, and data literacy. Thus, an ostensibly neutral AI systems will shape hierarchies of merit and employability. Viewed as an institutional extension rather than a pedagogical tool, AI risks reproducing a digitally skilled elite aligned with global capital along with digitally neglected masses unless countered by robust, state-led, rights-based regulation.

7 CONCLUSION

For AI to genuinely democratise education in India, an equity-first policy and implementation strategy is essential. This entails bridging the digital divide through universal device provision, robust internet connectivity in rural and underserved areas, and multilingual AI platforms accessible offline. Public–private partnerships can be leveraged, but with strong regulatory frameworks to ensure affordability, data privacy, and algorithmic transparency. AI deployments should be accompanied by continuous teacher professional development, enabling educators to harness AI insights

while maintaining human-centred pedagogy. Importantly, evaluation metrics for AI in education must go beyond efficiency or test score gains to include indicators of inclusivity, cultural relevance, and reduction in achievement gaps across socio-economic groups.

If implemented with these safeguards, AI can indeed serve as a leveller—scaling personalised, high-quality learning to all children regardless of location or socio-economic status. But without systemic investment in infrastructure, ethical governance, and capacity building, AI risks becoming another technological layer reinforcing the inherent exclusivity of quality education in India's stratified system. The lesson from both empirical interventions and pandemic-era digital inequalities is clear: technology alone cannot disrupt entrenched disparities; it must be embedded in a rights-based, inclusive educational reform agenda that addresses the material and social conditions of learning. In this light, AI is best seen not as a stand-alone solution but as one component in a broader ecosystem of reforms aimed at building an equitable, resilient, and future-ready Indian education system.

REFERENCES

- Adam, W. (1868) *Adam's Reports on Vernacular Education in Bengal and Behar*. Calcutta: Home Secretariat Press. [↗](#)
- Alatas, A. (2021) 'Islamic educational institutions in the past,' *IIUM Journal of Religion and Civilisational Studies*, 4(2), pp. 192–199. [↗](#)
- Anagol P. (2006) *The Emergence of Feminism in India, 1850-1920*. London: Routledge.
- Annual Status of Education Report (ASER). (2024) *Beyond Basics*. Available at: <https://asercentre.org/wp-content/uploads/2022/12/ASER-2023-Report-1.pdf> (Accessed: 10 August 2025). [↗](#)
- ASER (2021) *Did Technology Enable Inclusive Education in Rural Indian in a Year of Lockdown?* Available at: https://img.asercentre.org/graphics/aserdigitalcheck2020fortufts_01.09.21.pdf (Accessed: 10 August 2025). [↗](#)

- Bellenoit, H. (2012) 'Paper pens and power between empires in North India, 1750-1850,' *South Asian History and Culture*, 3(3), pp. 348–372. [↵](#)
- Bhabha, H. (1984) 'Of mimicry and man: The ambivalence of colonial discourse,' *Discipleship: A Special Issue on Psychoanalysis*, 28, pp. 125–133. [↵](#)
- Bhattacharya, S. eds. (2002) *Education and the Disprivileged: Nineteenth and Twentieth Century India*. New Delhi: Orient Longman.
- Buchanan, F. (1812) *An Account of the Districts of Bihar and Patna in 1811-1812*. Patna: Calcutta Oriental Press. [↵](#)
- Cheney, G.R., Ruzzi, B.B. and Muralidharan, K. (2005) *A Profile of the Indian Education System*. Washington, DC: National Center on Education and the Economy.
- Dharampal (2000) *The Beautiful Tree: Indigenous Indian Education in the Eighteenth Century*. New Delhi: Biblia Impex.
- Dundar, H. et al.. (2014) *Student Learning in South Asia: Challenges, Opportunities, and Policy Priorities*. Washington DC: World Bank.
- Global Information Society Watch (2019) *Artificial Intelligence: Human Rights, Social Justice and Development*. Available at: <https://www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development>. (Accessed: 13 August 2025).
- GOI (1992) *Programme on Action*. Available at: https://www.education.gov.in/sites/upload_files/mhrd/files/upload_document/npe.pdf (Accessed: 10 August 2025). [↵](#)
- GOI (2002) *The Constitution (Eighty-sixth Amendment) Act, 2002*. Available at: <https://www.india.gov.in/my-government/constitution-india/amendments/constitution-india-eighty-sixth-amendment-act-2002> (Accessed: 10 August 2025). [↵](#)
- GOI (2020) *National Education Policy 2020*. Available at: https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf (Accessed: 10 August 2025). [↵](#)
- GOI (2024) *Report on Unified District Information System for Education Plus (UDISE+)*. Available at: https://www.education.gov.in/sites/upload_files/mhrd/files/statistics-new/udise_report_existing_23_24.pdf (Accessed: 10 August 2025). [↵](#)

- Government of India (GOI) (1968) *National Policy on Education, 1968*. Available at: https://www.education.gov.in/sites/upload_files/mhrd/files/document-reports/NPE-1968.pdf (Accessed: 10 August 2025).
- Grant, C. (1792) *Observations on the State of Society among the Asiatic Subjects of Great-Britain, Particularly with Respect to Morals; and on the Means of Improving It*. London: House of Commons.
- Hossain, M.S. (2024) 'Exploring indigenous education in colonial india: A Historical Analysis,' *NSOU Journal of Social Sciences*, (1), pp.8–19. [↵](#)
- Hunter, W. (1882) *Report of the Indian Education Commission*. Calcutta: Superintendent of Government Printing. [↵](#)
- Jaiswal, A. and Arun, C.J. (2021) 'Potential of artificial intelligence for transformation of the education system in India,' *IJEDICT*, 17(1), pp. 142–158.
- Jana, T. and Sarkar S. (2021) 'A nation within a Nation: English education as a tool of divide and rule policy in colonial India,' *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 13(1), pp. 1–7. [↵](#)
- Khan, A.N.. (1983) 'The education system in muslim India: A historical perspective,' *Paedagogica Historica*, 23(1), pp. 67–83. [↵](#)
- Kincheloe, J. (2006) 'Critical ontology and indigenous ways of being: Forging a post-colonial curriculum,' in Y. Kanu (ed.) *Curriculum as Cultural Practice*. Toronto: University of Toronto Press, pp. 181–202. [↵](#)
- Kothari, D.S. (1966) *Report of The Education Commission 1964-66*. New Delhi: Government of India. [↵](#)
- Kumar, K. (2014) *Politics of Education in Colonial India*. New Delhi: Routledge.
- Lin, C.C., Huang, A.Y.Q. and Lu, O.H.T. (2023) 'Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review,' *Smart Learning Environment*, 10(41), pp. 1–22. [↵](#)
- Lopez, S. *et al.*. (2024) 'Artificial intelligence challenges and role of sustainable education in india: problems and prospects,' *Library Progress International*, 44,(3), pp. 18261–18271. [↵](#)
- Macaulay, T.B. (1835). *Macaulay's Minute on Education*. Calcutta: C.B. Lewis. [↵](#)
- Maitra, A. (2023) 'Becoming Sahibs: Bengali Bhadrlok travel cultures and a colony in paschim, c. 1850–1911,' *Language, Literature, and*

- Interdisciplinary Studies*, 5(3), pp. 1–24. [↗](#)
- Misra, S.N. and Ghadai, S.K. (2015) ‘Make in india and challenges before education policy,’ *Journal of Education and Practice*, (6), 97–102. [↗](#)
- Mudaliar, A.L. (1953) Report of the Secondary Education Commission (October 1952 – June 1953). New Delhi: Government of India. [↗](#)
- Muralidharan K. and Kremer, M. (2008) ‘Public-private schools in rural india,’ in R. Chakrabarti and P. Peterson (eds.) *School Choice Internation: Exploring Public-Private Parternships*. Cambridge: MIT Press, pp. 89–110. [↗](#)
- Naik, J.P. and Nurullah, S. (1974) *A Students’ History of Education in India (1800-1973)*. sixth revised edn. Delhi: The Macmillan Company of India Limited. [↗](#)
- National Policy on Education – 1986. New Delhi: Ministry of Human Resource Development, Government of India. Available at: https://ncert.nic.in/pdf/nep/Policy_1986_eng.pdf (Accessed: 10 August 2025).
- Nawani, D. and Jain, M. (2010) ‘Learners and learning in India: History, perspectives, and contexts,’ in Y. Zhao *et al.* (ed.) *Handbook of Asian Education: A Cultural Perspective*. New York: Routledge., pp. 503–528. [↗](#)
- O’Hanlon, R. (1994) *A Comparison between Women and Men: Tarabai Shinde and the Critique of Gender Relations in Colonial India*. Madras: OUP.
- PIB (2003) *Literacy in India: Steady March over the Years*. Available at: <https://archive.pib.gov.in/archive/releases98/lyr2003/rsep2003/06092003/r060920031.html> (Accessed: 10 August 2025). [↗](#)
- Radhakrishnan, S. (1949) *The Report of the University Education Commission (December 1948 – August 1949)* Volume I. New Delhi: Government of India. [↗](#)
- Rao, P.V. (2019) ‘Colonial state as ‘New Manu’? Explorations in education policies in relation to dalit and low-caste education in the nineteenth-century India,’ *Contemporary Education Dialogue*, 16(1), pp. 84–107.
- Roy, M. (2020) ‘AI intervention in education system of India: An analysis,’ *Solid State Technology*, 63(2), pp. 1395–1402. [↗](#)
- Sajid, M. and Hussain, M. (2025) ‘Educational policies and colonial legacy a historical analysis of successes and failures (1835-1947),’ *Global*

- Political Review*, X(I), pp. 75–88. [↗](#)
- Schiff, D. (2021) ‘Education for AI, not AI for education: The role of education and ethics in national AI Policy Strategies,’ *IJAIE*, (32), pp. 527–563. [↗](#)
- Shalini, and Tewari, A. (2020) ‘Sustainable education in india through artificial intelligence: challenges and opportunities,’ *WebSci’20 Companion*. Southampton, United Kingdom, July 6–10, Association for Computing Machinery. [↗](#)
- Srinivasan, V. and Murthy, H. (2021) ‘Improving reading and comprehension in K-12: Evidence from a large-scale AI Technology intervention in India,’ *Computer and Education: Artificial Intelligence*, (2), pp.1–22. [↗](#)
- Suman, A.K. and Shanu, S.K. (2024) ‘Exploring hindu college calcutta: Catalyst of intellectual evolution and its detractors,’ *History of Education*, 53(5), pp. 876–892. [↗](#)
- Velaskar, P. (2010) ‘Quality and inequality in indian education: Some critical policy concerns,’ *Contemporary Education Dialogue*, 7(1), pp. 58–93. [↗](#)
- Venkatarayanan, S. (2015) ‘Economic liberalization in 1991 and its impact on elementary education in India,’ *SAGE Open*, 5(2), pp. 1–13.
- World Bank (No Date) *District Primary Education Project*. Available at: <https://projects.worldbank.org/en/projects-operations/project-detail/P010464> (Accessed: 10 August 2025).
- World Economic Forum (2024) *Shaping the Future of Learning: The Role of AI in Education 4.0*. Available at: https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Learning_2024.pdf (Accessed: 13 August 2025). [↗](#)
- Zastoupil, L. and Moir, M. eds. (2013) *The Great Indian Education Debate: Documents Relating to the Orientalist-Anglicist Controversy, 1781-1843*. London: Routledge. [↗](#)

Ancient Wisdom for an Artificial Intelligence Age: Bridging Classical Ethics and Technological Futures

Sumit Kumar and Shashi Singh

University of Delhi

DOI: [10.1201/9781042018116-16](https://doi.org/10.1201/9781042018116-16)

ABSTRACT: In the age of rapid development of artificial Intelligence, modernity displays an incongruous relationship between rapid technological advances and the traditional preservation of moral standards, which retain deep roots of classical traditions. The following research examines how some of the most fundamental themes of ancient Indian and Greek traditions of thought, which deal with virtue, ethical restraint, balance, and a sense of community responsibility, may inform and influence modern digital progress. Using a comparative, interdisciplinary method, the exploration combines philosophical postulations in classical writings with contemporary analyses of AI ethics. An in-depth study of how technological systems with a prediction, control, and optimisation orientation often marginalise the concerns of meaning, purpose, and depth of relations is conducted. By a blend of historical contemplation and ethical analysis, the paper will argue that classical wisdom offers essential materials to rethink AI technology. The results show that the appeal to ancient ethical traditions can help to develop a more humanistic and contemplative technological ecosystem, an ecosystem that will elevate humility, connectedness, and self-restriction. This offering is especially relevant to researchers, techno-experts, and policymakers involved in designing environment-friendly, ethical projections of artificial Intelligence.

Keywords: Ancient Ethics, Artificial Intelligence, Technological Modernity

1 METHODOLOGY

This research paper adopts a qualitative, comparative, and interdisciplinary methodology that integrates classical philosophical analysis with contemporary ethical inquiry into artificial intelligence. Primary texts from ancient Indian and Greek ethical traditions are examined alongside modern AI ethics literature to identify convergences and tensions in concepts of virtue, responsibility, and restraint. Through interpretive analysis and normative reasoning, the study evaluates how classical ethical frameworks can inform more human-centered and reflective technological futures.

2 INTRODUCTION

Modern society is at a pivotal point, with the rapid development of artificial Intelligence (AI) and digital technologies that are quickly redefining the way people think, communicate, work, live, and govern. With intelligent machines increasingly invading our lives, some anticipate a future in which they make decisions in justice and healthcare systems. In contrast, others are convinced that they can simulate creativity through the use of generative models. The question is: what kind of future are we producing, and whose? Although AI offers efficiency, speed, and improved problem-solving, it has also raised serious ethical concerns regarding agency and biases, accountability, and the potential destruction of human values ([Boddington, 2017](#); [Crawford, 2021](#)). This paper addresses the ascendant structural misalignment between the instrumental rationality of artificial intelligence, which assigns primacy to efficiency and prediction, while relegating moral concerns to an ancillary position, and the ancient ethical traditions, where knowledge is inextricably linked with responsibility and the greater pursuit of human flourishing.

This misalignment has become a canonical problem of modernity, with technological progress on parade, lacking any moral underpinning. Being, first of all, the logic of control, prediction, and optimisation ([Zuboff, 2019](#)), AI technologies often operate without broader considerations of meaning, purpose, or ethical responsibility. It is in the light of this perspective that the renewed reliance on the ancient wisdom traditions of other civilisations, including, above all, India and Greece of the ancient world, can provide such an effective tool for reexamining the path of our technological present and future. They focused on ethical thinking, self-restraint, and the development of virtue at the centre of human existence, with the conception of knowledge as incapable of being separated from responsibility and flourishing (eudaimonia or Dharma) ([Nussbaum, 2001](#); [Ganeri, 2014](#)).

The differences between ancient Greek and Indian philosophies are riveting in terms of their merits, self-control, social beliefs, and the concept of the good life. Whereas Greek philosophers, such as Socrates, Plato, and Aristotle discussed the role of logos, phronesis (practical wisdom), and arete (virtue) in the construction of a just society, the Indian philosophical systems, including the Upanishadic, Buddhist, and classical yoga traditions, designated ideals of Dharma (ethical duty), ahimsa (non-violence), and niyama (self-discipline) as paths towards a self-realisation and social peace ([Olivelle, 1998](#); Bilimoria). Both schools warned against the mindless pursuit of power and emphasised the need for inner transformation as a prerequisite for worldly wisdom.

This paper suggests that the wisdom imparted by these classical traditions provides valuable resources for addressing ethical issues related to AI and technological governance. Filling the gap between ancient and current perspectives in AI ethics and digital design, the paper aims to develop an image of technology that values not only innovation but also human flourishing, moral restraint, and communal well-being ([Floridi et al., 2018](#)). In doing so, it does not promote a nostalgic, backwards look at the past, but rather a constructive movement of historic ethical systems towards contemporary techno-social existence.

Therefore, the main object of our investigation, which is, in the first place, to dig out and analyse fundamental ideas in Greek and Indian ethics; in the second place, to discuss their application to the construction of a more philosophic and human future of technology. This interdisciplinary

intervention contributes to a growing body of literature that critiques the instrumental reasoning of AI and redefines digital modernity as more than just an arena of technical benefits, but also one of moral soundness and philosophical creativity.

3 ANCIENT GREEK WISDOM AND ARTIFICIAL INTELLIGENCE

The rapid development of the field of Artificial Intelligence (AI) has led to drastic changes in the way societies work, make decisions, and regulate people's lives, both consciously and subconsciously. However, this pace of technological development has raised philosophical and ethical issues related to autonomy, responsibility, and what constitutes human flourishing. Since machine learning and algorithmic systems are rapidly becoming a filter through which human existence is mediated, the wisdom traditions of the classical civilisations, most importantly that of ancient Greece, provide an antidote to the utilitarian and efficiency-driven rationality that characterises modern technological culture. Ancient Greek philosophy (with its focus on reason, virtue, and telos/purpose) speaks to indifferent contemporary perspectives on AI and digital modernity and opens up more ethical perspectives on these issues (Broadie, 2007.).

3.1 *Greek ethical philosophy: Reason, virtue, and the good life*

The main tradition of Greek ethical thought, which developed during the 5th-4th centuries BCE, has been directed at eudaimonic cultivation, which means more than being happy, as it is a state of flourishing based on virtue and intellectual excellence ([Cooper, 1986](#); Kraut, 2018). This pursuit cannot be reduced to either subjective fulfilment or material prosperity; it is instead a totality of being obtained by coordinating action, character, and rational abilities with the advantage of the individual and the society.

To Socrates, eudaimonia could never be discussed outside the context of the examined life, characterised by constant ethical questioning and a search for self-knowledge (Plato, Apology 38a). He argued that virtue

(arete) is a kind of knowledge, and that moral weakness is the result of ignorance, not vice ([Brickhouse and Smith, 2004](#)). According to this epistemic view of ethics, intellectual engagement is central to a life of morals, which means that the quest for knowledge and the development of moral character mutually support each other ([Vlastos, 1991](#)). Socratic ethics, therefore, does not allow a firm distinction between knowing and doing the Good.

Plato expanded the teachings of his teacher and anchored the ethics in a metaphysical and political project. In *The Republic*, he introduced justice as a kind of harmony, both in the soul and the state, where reason should dominate and guide spiritedness and appetite, allowing each element to play its respective role (Plato, *Republic* 435b-441c; Reeve, 1988). The axis of this vision is the form of the Good, the timeless, unchanging criterion that informs all the other values (Plato, *Republic* 508e509b). This structure gives rise to the concept of the philosopher-king: any other ruler cannot rule rightly, as they lack not only theoretical insights into the Good but also the moral leadership necessary to guide communal life into a flourishing state (Annas, 1999). This perception encompasses the notion of political legitimacy in conjunction with moral and intellectual excellence, emphasising the subordination of power to wisdom.

Although Aristotle was a debtor to his predecessors, he captured attention by moving away from metaphysical ideals and turning to the practice of human beings. In the *Nicomachean Ethics*, he outlined virtue as a steady inclination to do and feel the things that are considered as the “mean” between excess and deficiency, identified by reasoning and responsive to the specifics of the case at hand (Aristotle, *Nicomachean Ethics* 1106b1140b; Hursthouse and Pettigrove, 2022). Moral excellence, in the case of Aristotle, is developed through habituation, a lifelong process of internalising the desires, emotions, and reasoning powers towards noble things (kalon). His concept of practical wisdom (phronēsis) holds a pivotal position in his ethics: it is the intellectual virtue of theorising about the means to make a good life a reality by incorporating the imperatives of received wisdom with the contingencies of individual experience (Burnyeat et al., 1980).

Together, these theorists oppose the reductionism of Intelligence as being either a technical or procedural skill. In the Greek tradition, logos (reason)

does not involve abstract computational capacity, but rather an ethically oriented faculty of community responsibility ([Nussbaum, 2001](#); [Sherman, 1997](#)). In this regard, the ethics in Greek may serve as a penetrating critique of any form of thought system (ancient or otherwise) that separates knowledge and ethics, or strength and moral responsibility, thereby linking Ancient Greek wisdom with Artificial Intelligence.

3.2 *Ancient greek ethics and the moral limits of artificial intelligence*

The conversation on innovative artificial Intelligence and ancient ethical traditions revolves around a crucial distinction: the distinction between the capability to use information, which is Intelligence, and the ability to reach morally adequate decisions, also known as wisdom. Although AI systems can replicate specific cognitive processes, including learning, linguistic generation, and pattern recognition, they are incapable of conscious awareness, moral agency, and contextual sensitivity, essential components for an ethical life ([Floridi et al., 2018](#); [Coeckelbergh, 2020](#)). This distinction was keenly felt by ancient Greek philosophers, particularly in the context of describing *phronēsis* (practical wisdom). To Aristotle, *phronēsis* referred to the capacity through which humans can ponder what is good and expedient in the realm of a whole life, thereby bringing together the sentiments of understanding and making rational judgments (Nicomachean Ethics 1140a-1140 b; [Sherman, 1997](#)). Such are the very features missing in algorithm systems that only make decisions constrained by the parameters of their programming and the biases of their training data.

It is often said that the logic of operation of AI can be explained through control, prediction, and optimisation, which is a sort of algorithmic rationality ([Zuboff, 2019](#); [Mittelstadt et al., 2016](#)). This reason prizes expediency and correctness above ethical subtlety and social welfare, which Greek philosophers would have considered technology to have been uncoupled from ethical governance. In Plato, the *Gorgias* (456d 457c), the craft and the technology are morally indifferent means, which can have moral relevance only when they are put in the second priority to the search for the Good. The same word of caution was reiterated by Aristotle (*Politics*

(1260a)), who cautioned against knowledge that has not been united with virtue and which would be considered incomplete, bringing even harm to the possessor of the knowledge. In the era of AI-based governance, automated decision-making, and predictive analytics, such issues take on a new sense of urgency, as eliminating moral reasoning in these situations may lead to technically effective but morally harmful outcomes.

The reconsideration of these classical insights presents an opportunity to reevaluate AI ethics in a manner that does not imply narrow models of compliance. According to the Greek perspective, technology needs to be governed not by the question of what can be done with machines, but by the question of what needs to be done with machines in terms of facilitating human flourishing. This reframing aligns with more contemporary requirements that both ethical thinking and the well-being of relationships be considered in the drafting and application of AI ([Boddington, 2017](#); [Crawford, 2021](#); [Floridi et al., 2018](#)).

With the redistribution of lines of modernity facilitated by AI, the danger of subordinating a form of moral deliberation to technological innovations is becoming increasingly urgent. The uncritical implementation of predictive algorithms tends to undermine human abilities to act, empathise, and discern morally, as they are replaced by a black box of unexplainable decisions, which evade societal scrutiny (O Neil, 2016; [Pasquale, 2015](#)). Making the situation worse, AI systems, however sophisticated they may be, lack the moral and spiritual aspects that have guided humans throughout millennia.

When facing it, an ancient Greek morality compels us not to think about AI governance solely in a utilitarian or individualistic way. Instead, it leads to relationality, shared responsibility, and the development of virtues that protect humans and their dignity. The fact that *phronēsis* must be included in AI discussions does not mean that machines will be able to attain moral awareness or that they should; instead, it requires human designers and policymakers to be responsible for instilling moral sagacity into the frames and functions of technological systems. In such a way, the gap between the blistering digital advancement and the eternal human pursuit of a better life and worth living is bridged.

The inclusion of ancient ethical reasoning in the debate on artificial Intelligence can also be seen as problematic in the sense that it highlights

the need to reconsider some concept of epistemic authority in the era of machine decision-making. This trend was especially prominent within Greek philosophy, particularly in the Aristotelian and Platonic traditions, where great stress was placed on the development of intellectual virtues, as well as sophia (theoretical wisdom) and episteme (knowledge of first principles), in addition to phronēsis. The virtues could not and were not separated, especially with moral development, as it was a compound ideal in which there was an interdependence of knowing well and living well ([Annas, 1993](#); [Cooper, 1986](#)). Comparatively, modern AI systems act in the loosely circumscribed realms of epistemic value, that is, they are engineered to maximise performance in terms of specific, functional criteria without necessarily directing the results towards any understanding of the overall Good. This generates an ethical asymmetry: ethical responsibility is assigned to human agents, while a system of their own making, algorithmic agents, can have major social implications without having the ethical responsibility or ability to correct their work in the context of ethical analysis ([Johnson, 2015](#); [Bryson, 2018](#)). A pertinent example comes from healthcare, where a widely used risk-prediction algorithm in the United States was found to disadvantage Black patients. The tool was designed to identify which patients needed to receive extra care management resources, but instead of directly assessing medical need, it relied on healthcare costs as a convenient proxy. Due to widely rampant systemic inequality, less money is spent on the treatment of black patients in comparison to white patients. The implications were such that, despite being equally sick as white patients, black patients were ranked as lower risk. Consequently, many black patients were denied the adequate healthcare they needed. This is precisely the kind of rigid, context-blind reasoning that the Greek philosophers warned against. An approach informed by phronesis can help rectify this bias (Vartan, 2019).

Such a delegation to a morally neutral body as a consequence of such decision-making would be understood by the Greek tradition as an absolute failure of political and ethical judgment, as power without virtue is inimical to the validity of any ruling. Applied to the case of AI, this implies that the technology will have to be institutionally and structurally bound to practices and institutions that can discuss and deliberate about the goals of action, so

that innovations can be directed toward the prosperity of a polis, rather than the exigencies of market efficiency or state control over them.

4 ANCIENT INDIAN VALUES AND THE ETHICS OF AI

Ancient Indian philosophy provides unique perspectives on reality, ethics, and knowledge in various matters, encompassing a broad range of issues. It encompasses an intense and enormously diverse intellectual background. Ancient Indian thought is typically comprehensive and is integrated with metaphysics, epistemology, ethics, and religious practice, which differs radically from the hermetically sealed isolation of Western philosophy. (Datta and Chatterjee, 1984).

Essentially, ancient Indian philosophy is centred on the interdependence of all beings, the pursuit of liberation (moksha) from ignorance and suffering, and the cultivation of virtues (Dharma) and wisdom (jñāna). It comprises both heterodox schools (Nāstika), such as Buddhism, Jainism, and Cārvāka, which challenge or dispute Vedic authority, as well as orthodox schools (Āstika), including Vedānta, Nyāya, Sāṅkhya, Yoga, Mīmāṃsā, and Vaiśeṣika, which accept the authority of the Vedas. The foundation of spiritual and moral consideration in such schools comprises central ideas such as karma (the principle of action and consequence), ahimsā (non-violence), ātman (the self), and brahman (world consciousness). For example, Nyāya schools emphasise epistemic duty and provide elaborate critiques of perception, reason, and inference.

Ancient Indian philosophy foregrounds moral interdependence, spiritual intentionality, and the integration of knowledge and being, contrasting with the prevailing Western model, which prioritises individual liberty, calculative rationality, and legalistic moral reasoning. Ancient Indian philosophy is therefore well-positioned to provide alternative models for assessing the ethical implications of emerging technologies such as artificial Intelligence ([Vallor, 2016](#)).

4.1 *Dharma: Ethical duty in a contextual world*

One of the fundamental concepts in ancient Indian philosophy is Dharma, which encompasses the right duty, law, and cosmic order. Unlike West's codified notions of law or ethics, which exhibit a deontological strictness, the Indian concept of Dharma displays considerable relativism. According to the Bhagavad Gītā, the concept of Dharma primarily consists of two key notions: svadharma (personal duty) and nishkāma karma (selfless action).

Grounded on this basis, a study attempts to lay bare the ethical nuances of the automated machines. Arguing that Western ethical theories, in their liberal, universalist prescriptions, do not capture the social and cultural diversity of India, the authors take their stance on 20 genuine ADM use-cases to reveal how automated reasoning is set to conflict with the pluralistic nature of Indian society. Alternatively, they suggest embracing frameworks based on Dharma, Nyāya, and Karma to navigate the uncertainties fueled by the rise of artificial Intelligence. ([Biju and Gayathri, 2025](#))

This context-sensitive model is particularly pertinent in domains such as welfare and medicine, where algorithmic recommendation has tangible ramifications. The universal AI reasoning regime dismisses relational duties inherent in Indian ethics. A Dharma-guided methodology, solidly rooted in svadharma, permits exceptional situations to be taken into consideration instead of generalisable rules, in contrast to Western ADM systems' adherence to one-size-fits-all approaches.

Similarly, nishkāma karma—action performed for its own sake, without attachment to the results—offers a critique of profit-driven or predictive control AI. It favours intention and moral purpose over efficiency. Combined with lokasaṅgraha (the goodness of the world), it can shape AI as a public service tool, rather than one of private domination, insisting on designs based on pluralistic values.

4.2 *Nyāya: Valid knowledge and epistemic responsibility*

In addition to these ethical frameworks, we might draw inspiration from the ancient Indian school of Nyāya, not merely as a logic, but as an epistemic tradition rooted in public reason. Nyāya places great emphasis on pramāṇa, the reliable sources of knowledge, specifically anumāna (inference) and

vāda (formal dispute), as well as procedural justice (nyāya in the sense of fair adjudication). Current scholarship has made evident how Nyāya's answer to the Cārvāka objection to inference anticipates modern solutions to the “problem of induction,” providing a systematic response to algorithmic obscurity ([Mukhoty, 2024](#)). In AI ethics, an approach inspired by Nyāya would prioritise dialogical reasoning, epistemological transparency, and the accountability of the process of making knowledge claims.

In contrast to instrumental rationality and black-box determinism of much Western automated decision-making (ADM) technology, ancient Indian ethics are concerned with justifiability and interpretability.

To add more depth to our argument, we could also engage with the concept of Ānanda. Though it is traditionally interpreted as bliss, it carries a much deeper meaning. It is a state of thriving in all spheres: integrative, spiritual, and social. In the Taittirīya Upaniṣad, Ānanda is the supreme state of human thriving, which transcends pleasure and knowledge, and is deeply related to good living and social harmony. Ānanda can potentially alter the telos of technological development: it could shift AI's goal from profit maximisation to lokasaṅgraha. Such a multifaceted approach aligns with Sri Aurobindo's vision of spiritualised reason, a form of cognition that does not rely blindly on instrumental logic. This understanding transcends instrumental logic to incorporate ethical, intuitive, and other relational aspects of Intelligence that are otherwise glossed over. With this strategy, AI will not only be used as an instrument of control; it will instead support human thriving by prioritising values that are epistemically deep and morally nuanced.

4.3 *Niyama and yama: The disciplines of inner and outer ethics*

Likewise, the Yoga Sūtra of Patañjali provide a two-pronged approach to answering the question of ethics: the yamas (social restraints) and niyamas (personal disciplines). These comprise values such as satya (truthfulness), aparigraha (non-possessiveness), santosha (contentment), and svādhyāya (self-study). In a world incessantly dominated by surveillance capitalism and its surrounding data accumulation and digital pervasiveness, the

importance of these fundamental principles of ancient Indian philosophy is as vital as ever. The Yamas and Niyamas facilitate the evolution of intentionality at both the personal and systemic levels. The AI systems developed on the principles of aparigraha would limit data collection to what is necessary alone. They would minimise intrusive personalisation and not monetise attention. Likewise, svādhyāya can also introduce transparency and accountability in AI systems, not only in how they are written but also in assumptions and intentions with which they act.

This moral framework also applies to individuals who create and deploy AI. Developers and engineers would be willing to approach their work as sādhanā (spiritual practice), where technical perfection is inseparable from moral clarity.

4.4 *Karuṇā and maitrī: Compassion and friendship in the algorithmic age*

Besides codified rules and principles, Indian traditions most insistently promote the cultivation of virtues and emotional dispositions that support ethical life. Among the virtues are compassion (karuṇā) and friendship or loving-kindness (maitrī). These are not passive sentimentalities, but active commitments to the well-being of others grounded in a sense of interdependence. These virtues are grounded in relational contexts that esteem responsiveness over adherence to rules. ([Heim, 2015](#)) [Harvey \(2000\)](#) holds that in the Buddhist tradition, karuṇā and maitrī are cultivated through deliberate practices that transform perception and action toward others.

In artificial Intelligence, the possibility of personalisation turning into manipulation and prediction invading the agency of individuals offers another vision of the relationship between humans and technology. What are the implications of the idea of AI systems being designed to emphasise compassion over control? How do we create platforms that engender solidarity rather than fragmentation, dialogue rather than echo chambers? This moral turn is echoed in the work of [Hongladarom \(2020\)](#), who argues that Buddhist compassion can be employed to create data ethics by

incorporating relational responsibility and compassion rather than abstract rights or rules.

Empathic artificial Intelligence is not merely a utopian dream; it could also be a feasible development based on deliberate design principles. However, making it a reality would require constant effort, and it will depend on interdisciplinarity and intercultural cooperation, as well as a willingness to ensure that ethical considerations are placed at the centre of technological development. Some recent studies have advocated a return to virtue ethics in the development of emerging technologies, thereby emphasising that the development of moral character — both individual and institutional, is necessary to create technologies that enable human flourishing rather than its decline. ([Vallor, 2016](#))

4.5 *Advaita and AI: Consciousness in code*

Another dimension that Indian philosophical schools of thought contribute to AI ethics is the concept of non-duality (advaita), which posits that everything is interconnected. Advaita Vedānta holds that all differences, between the self and the other, the mind and the body, and the human and the non-human are ultimately illusory. The self (ātman) is not separate from the world (brahman); instead, it is essentially one with it. This metaphysical doctrine has considerable ethical import: it promotes a sense of shared responsibility, where causing harm to another is, in a more profound sense, causing damage to oneself. This worldview challenges the distinction between designer and user, as well as between system and society, within the realm of artificial Intelligence.

Rather than limiting our inquiry to what artificial Intelligence might accomplish for isolated individuals, a more expansive framework urges us to consider the forms of mutuality and ethical integration possible across every layer of technological engagement. An AI based upon the philosophy of non-duality would not view data subjects as objects to be manipulated, but as co-constitutive coparticipants within a shared ontological space. This view emphasises co-presence rather than unilateral control, aligning with larger ecological and social systems, and reconceives Intelligence not in

terms of dominion over complexity, but as an aestheticised sense or sensitivity to it.

In this context, Advaita presents a fundamental challenge to metaphysical assumptions in conventional AI thinking. It calls for new paradigms of design that prioritise interdependence as the ethical foundation of technological progress. The moral dimensions of Advaita are not abstract, but also enacted in practices of compassion, humility, and restraint that follow spontaneously from the experience of oneness. Such a metaphysics transforms our conception of AI: it is not an external tool to be maximised, but an extension of the human ethical field in relation.

One of the most critical dimensions of modern AI ethics that has thus far been underexplored is the place of intention and moral psychology. While dominant practices are short-sighted in their ends-prioritisation and deontological respect for rules, ancient Indian philosophy focuses on cultivating moral agency. These are the sharpening of intention (*saṅkalpa*), non-attachment to egoistic desire (*niḥsvārtha*), and constant self-reflection (*svādhyāya*).

In the Bhagavad Gītā, action is not just limited to its fruit or consequence; action is instead assessed following the intention of the doer and whether it conforms to Dharma or not. The concept of *niskāma karma*, action without attachment to results, emphasises refining the motive rather than merely maximising profit. This approach to ethics has the potential to critique the utilitarian bias inherent in AI systems. In such systems, efficiency is privileged at the expense of intentional horizons of human agency.

The Yogic and Buddhist traditions build on this with particular practices for the development of consciousness. In Patañjali's Yoga Sūtras, the eightfold path of *aṣṭāṅga yoga* progresses to *samādhi*, a state of inner concentration characterised by moral clarity and non-egoic consciousness. The *niyamas*, for example, *śauca* (purity), *tapas* (discipline), and *īśvarapraṇidhāna* (surrender to a higher order), are not discursive principles, but inner practices intended to generate moral disposition from the inside out.

From this perspective, ethical action is not an add-on to technical competence, but a manner of being, cultivated through mindfulness practice, self-discipline, and humility. These inner leanings—*dayā*

(compassion), kṣamā (forgiveness), and maitrī (loving-kindness), cannot be directly translated into outcomes, but rather serve as the affective underpinnings of an ethical community ([Heim, 2015](#); [Harvey, 2000](#)).

When applied to artificial intelligence, a Dharma-inspired approach would mean that AI systems will be assessed not only for technical efficiency but also on the grounds of their alignment with justice, ecological responsibility, and respect for human dignity. For example, the COMPAS system in the United States has been criticized for disproportionately flagging minority individuals as high risk. From an efficiency standpoint, this is workable; ethically, though, this involves abandoning the exercise of upholding social equities. A Dharmic approach would try to correct this through constant auditing of bias so that predictive accuracy is not maximized at the expense of fairness or human dignity. Similarly, a Nyaya-guided system would call for intelligibility of the decision-making process so that affected persons can challenge or understand decisions. For example, applicants are frequently provided with only a numeric score without disclosing reasons for refusal of a loan. A Nyaya-oriented system would demand grounds of decision to be articulated—be it repayment history, income stability, or other factors—so that the judgment can be measured in terms of canons of fairness and accuracy.

Using these concepts to address the ethics of artificial Intelligence is an issue of reframing questions from “What should AI do?” to “What sort of intentional structures and relational dispositions are built into AI systems and designers?” This is especially important in the age of affective computing, where computers are being designed to perceive and react to the emotions of human beings. Without careful reflection on intention and virtue, these technologies risk being used as instruments of manipulation instead of fostering mutuality. Secondly, Indian philosophies oppose the notion of a fixed, independent self which underlies much Western moral theory. Buddhist anātman and the emphasis of Advaita on non-duality challenge the concept of an independent, enduring agent. Moral agency within these philosophies is relationally negotiated and fluid, and it calls for systems of care and responsibility sensitive to impermanence, interdependence, and humility. Indian philosophical ethics cannot be separated from metaphysics; it arises from one's experience of dependent reality, and not only because of reason.

Embracing such values in AI design and policy would make possible systems that are less about extraction and more about reflexivity, less focused on domination and more sensitive to co-flourishing. Engineers, designers, and policymakers would need to perform their work as *sādhana*, wherein technical competence and moral intentionality complement each other. By highlighting the moral psychology of intention, the dynamic nature of selfhood, and the cultivation of inner virtue, ancient Indian philosophy offers a rich and deep counterpoint to prevailing theories of rational control, prediction, and optimisation of behaviour. Rather than merely adding to the prevailing models, ancient Indian values call for a fundamental rethinking of the moral framework on which AI systems are based.

The dilemmas that AI poses are not confined merely to engineering or policy concerns. They are essentially concerns pertaining to our collective humanity. To effectively address the issues, the Indian Value system offers a wealth of ethical concepts. By moving away from a profit-oriented, myopic worldview, our ethics encourage us to adhere to a more contemplative and constrained approach to technological advancement.

What is also needed is a fundamental overhaul of the concept of innovation. Hitherto, we have prioritised speed, dominance, and abstraction. However, innovation informed by Indic values will encourage more dialogue, discernment, and empathy. Such an approach does not entail compromising with the rigour of innovation; it merely advocates a more holistic and inclusive technical development.

Essentially, marrying modern technology with ancient ethical frameworks enables us to reexamine what it means to be human in an era of increasingly sophisticated artificial machines.

5 COMPARATIVE ETHICAL DIALOGUES: BEYOND GREECE AND INDIA

While the rich ethical traditions of Greece and India provide us with a robust foundation to navigate through the uncertainties stoked by the burgeoning of artificial intelligence, other civilizations, too, offer vital

resources that can enhance our understanding and help us move away from the instrumental and myopic rationality. In particular, Confucian ethics and African Ubuntu philosophy highlight several dimensions of communal well-being and relationality that are of paramount importance in this age of artificial intelligence.

Confucius created a humane ethical system, as opposed to a legal institution, to govern and stabilise the country by mobilising, rather than repressing people's emotions. Accordingly, he established a comprehensive ritual procedure (礼) for people to follow the formal rules, and instituted human ethical relations (人伦) as a basis for informal personal interactions. The Confucian value system is deeply rooted in notions of harmony and relationality. What this means in the context of our present discussion is that ethical evaluation cannot be evaluated merely on the basis of utilitarian calculations alone; it must account for the quality of human relationships and the broader social order in which we deploy artificial intelligence. For instance, the use of artificial intelligence in education would not be confined merely to efficiency; it should also include discussions on whether or not it succeeds in fostering a respectful teacher-student relationship.

Similarly, in the case of Ubuntu philosophy, we see that humans are situated in the broader context of the environment they inhabit. The maxim “umuntu ngumuntu ngabantu” perfectly encapsulates the essence of this community relationship: a person is a person through other persons. When extended to the pressing question of artificial intelligence, the Ubuntu philosophy emphasizes justice and inclusion. For instance, the management of AI systems in predictive policing, healthcare, or resource distribution cannot ethically be determined by technical precision alone. The philosophy calls for these systems to be measured by their ability to advance the general good of the community, by fairness, and by the protection of the vulnerable. By incorporating marginalized voices in the decision-making process, the philosophy would ensure that solutions to the most pressing challenges of our times are not predicated on detached and technocratic measures.

What is noteworthy here is the fact that ancient cultures across the globe — irrespective of geographical and cultural constraints — converge on the question of human flourishing, responsibility, and restraint, even if the lexicon in which these are articulated might differ. By incorporating these

traditions in the development of artificial intelligence, we can ensure that moral and ethical concerns are not made subservient to technical precision and efficiency.

6 CONCLUSION

The rapid development of artificial Intelligence has made the moral and philosophical issues surrounding it even more acute, necessitating that societies address the questions of the nature, scope, and purpose of technological advancements and their impact. Although modern theories tend to focus on the efficient and predictive aspects and relative competitiveness in the market, little has been said about the underlying ethical principles of human flourishing. In this respect, the Greek and Indian ethical traditions represent invaluable sources of inspiration that can be used to envision the relationship between people and machines.

The Greek approach to knowledge highlights the importance of developing *phronēsis* (practical wisdom), *arete* (virtue), and a commitment to *eudaimonia* (flourishing), as it is impossible to study and not appreciate the connection between this knowledge and moral responsibility. The same tradition cautions against the technical mastery pursued without regard for ethics and advises that power must be subservient to wisdom. The purpose of technology depends on human flourishing, not on control or expedience. In the context of AI, this perspective entails changes in the governance and design processes, shifting from asking what can be automated to what ought to be automated, and identifying the ethical limitations that should be considered along this path.

The Indian philosophical tradition develops and expands this vision by basing moral life on relationality, intention, and the unity of being. Ethics embodied by concepts like *Dharma*, *nishkāma karma*, *ahiinsa*, and *advaita* consists of an ethics that is circumstance-sensitive, non-dualistic, and service-oriented towards everyone. Unlike universalist and abstract rule-based morals, these schools acknowledge the context-sensitive nature of moral action, the epistemic responsibility, and the cultivation of inner virtues as a transformative moral practice. When they are applied in the

context of AI, they necessitate a transition from exploitative and abusive systems to ones designed for serving lokasaṅgraha, the collective well-being of all, and operating on the principles of compassion, transparency, and non-harm.

The fact that these traditions are being introduced into discussion does not imply a backwards-looking reverence for the past or a blind application of age-old conventions in a digital environment. Instead, it is a synthesis of creativity that cannot deny the relevance of moral self-cultivation, critical thinking, and responsibility in relationships among people in determining the future of technology. Such conventions do not acknowledge the reality of a divorce between knowledge and virtue, reminding us that the rectitude of power, political or algorithmic, is eventually evaluated by the excellence of its aims and the sagacity of its practice.

To establish these considerations in practice, the governance and design systems of AI would need to incorporate institutional frameworks that place a greater emphasis on moral deliberative processes than on technical innovation, and educational models that promote more ethically oriented education for engineers and computer scientists. The benefit of this would be that it allows for ensuring that AI does not become a black box and unaccountable force in human affairs, but rather a communicative and transparent participant in a shared ethical enterprise.

Just as we had a synthesis of ancient Greek and Indian wisdom at the end of part two, we are now allowed to enjoy a philosophical formulation of technology writ large. It does not frame AI as an automation tool without question, since it places that area of study as one where very core questions of human purpose, fairness, and well-being need to be addressed. By so doing, it calls upon us to repossess the moral imagination required to see that, when the intelligent machine age arrives, we will be not only clever but also sorely wise.

REFERENCES

Annas, J. (1993). *The Morality of Happiness*. Oxford: Oxford University Press. [!\[\]\(23d9fc146e83b5c3013cfa32c784f8d5_img.jpg\)](#)

- Aristotle, and Tredennick, H. (1976). *The Ethics of Aristotle : The Nicomachean Ethics* (; J.A.K. Thomson, Trans.; Rev. ed.). Penguin.
- Biju, P.R. and Gayathri, O., (2025). Indic approach to ethical AI in automated decision-making systems: Implications for social, cultural, and linguistic diversity in the native population. *AI and Society*. doi:[10.1007/s00146-025-02381-z](https://doi.org/10.1007/s00146-025-02381-z)
- Bilimoria, P. (2007). *The Indian Ethics: Classical Traditions and Contemporary Challenges*. London: Routledge
- Boddington, P. (2017). *Towards a Code of Ethics for Artificial Intelligence*. Cham: Springer.
- Brickhouse, Thomas C. and Smith, Nicholas D. (2004). *Routledge Philosophy Guidebook to Plato and the Trial of Socrates*. New York: Routledge. Edited by Nicholas D. Smith .
- Bryson, J.J. (2018). Patience is not a virtue: The design of intelligent systems and systems of ethics. *Ethics and Information Technology*, 20(1), pp. 15–26.
- Chatterjee, S.C. and Datta, D.M., (1984). *An Introduction to Indian Philosophy*. 8th edn ed. Calcutta: University of Calcutta.
- Coeckelbergh, Mark (2020). *AI Ethics*. Cambridge, Massachusetts, USA:
- Cooper, J.M. (1986). *Reason and Human Good in Aristotle*. Indianapolis: Hackett.
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., and Vayena, E. (2018). AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Ganeri, J. (2014). *The Lost Age of Reason: Philosophy in Early Modern India 1450–1700*. Oxford: Oxford University Press.
- Harvey, P., (2000). *An Introduction to Buddhist Ethics: Foundations, Values and Practice*. Cambridge: Cambridge University Press.
- Heim, M., (2015). *Theories of the Gift in South Asia: Hindu, Buddhist, and Jain Reflections on Dana*. London: Routledge.

- Hongladarom, S., (2020). *The Ethics of AI and Robotics: A Buddhist Viewpoint*. Lanham, MD: Lexington Books. [↗](#)
- Hursthouse, Rosalind and Pettigrove, Glen (2022). Virtue Ethics. Stanford Encyclopedia of Philosophy.
- Johnson, D.G. (2015). Technology with no human responsibility? *Journal of Business Ethics*, 127(4), pp. 707–715. [↗](#)
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., and Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2). <https://doi.org/10.1177/2053951716679679> (Original work published 2016). [↗](#)
- Mukhoty, A.P., (2024). Nyāya–Cārvāka Debate on Inference and the Problem of Induction. *Journal of Indian Philosophy*, 52(5), pp. 675–700. DOI: [10.1007/s10781-024-09585-5](https://doi.org/10.1007/s10781-024-09585-5) [↗](#)
- Nussbaum, M.C. (2001). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press. [↗](#)
- O'Neil, C. (2016). *Weapons of Math Destruction, How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Broadway Books.
- Olivelle, P. (1998). *The Early Upaniṣads: Annotated Text and Translation*. Oxford: Oxford University Press. [↗](#)
- Pasquale, F. (2015). *The Black Box Society, the Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/harvard.9780674736061>. [↗](#)
- Plato. (2000). *The Republic* (trans. G.M.A. Grube, rev. C.D.C. Reeve). Indianapolis: Hackett Publishing.
- Plato. (2002). *Five Dialogues* (trans. G.M.A. Grube). Indianapolis: Hackett Publishing.
- Schofield, Malcolm ; Burnyeat, Myles and Barnes, Jonathan (eds.) (1980). *Doubt and Dogmatism: Studies in Hellenistic Epistemology*. New York: Oxford University Press.
- Sherman, Nancy (1997). *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. New York. [↗](#)
- Vallor, S., (2016). *Technology and the Virtues: A philosophical Guide to a Future Worth Wanting*. Oxford: Oxford University Press. [↗](#)
- Varshney, K.R., (2023). *Decolonial AI Alignment: Openness, Viśeṣa-Dharma, and Including Excluded Knowledges*. *arXiv*. Available at:

arXiv:2309.05030

Vlastos, Gregory (1991). *Socrates: Ironist and Moral Philosopher*. Ithaca, N.Y.: [↵](#)

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. New York: PublicAffairs. [↵](#)

Digitalization, Structural Reforms, and Big Data Ecosystems: The Role of Artificial Intelligence in India's Remarkable Economic Stride in Recent Decades

Priti Rai and Shantnu Kumar

University of Delhi

DOI: [10.1201/9781042018116-17](https://doi.org/10.1201/9781042018116-17)

ABSTRACT: The study explores the role of Artificial Intelligence (AI) in supporting India's remarkable economic stride in recent decades, with a focus on digitalization, structural reforms, and the growth of data ecosystems. It argues that AI has become a structural enabler of economic modernization rather than a peripheral technological tool. AI-driven technologies have strengthened India's digital public infrastructure, including Aadhaar, the Unified Payments Interface (UPI), and the Goods and Services Tax Network (GSTN), enhancing financial inclusion, improving service delivery, and enabling real-time data integration across sectors. In the domain of structural reforms, AI applications in predictive analytics, fraud detection, and automated compliance have improved governance efficiency, optimized tax administration, and reinforced welfare targeting, thereby enhancing state capacity. Concurrently, the rapid expansion of data ecosystems—fueled by mobile connectivity, digital payments, and e-governance platforms—has created fertile ground for AI-enabled policymaking, innovation, and productivity gains across agriculture, healthcare, manufacturing, and logistics. The article emphasizes that these developments collectively highlight AI's transformative potential in shaping India's institutional, economic, and technological landscape. It concludes that AI is central to India's digital transformation and long-term growth trajectory, demonstrating its critical role in supporting inclusive, data-driven, and technology-enabled economic development.

Keywords: Artificial Intelligence, Digitalization, Structural Reforms, Data Ecosystems, and Economic Growth

1 METHODOLOGY

The paper employs a qualitative policy-analytic methodology combining secondary data analysis with institutional and sectoral case examination. It analyses official reports, economic

indicators, and digital public infrastructure frameworks to assess AI's role in digitalization and structural reforms. A political-economy lens is applied to evaluate AI's contribution to economic growth within India's evolving data ecosystems.

2 ARTIFICIAL INTELLIGENCE AND INDIA'S DIGITALISATION – IMPLICATIONS FOR ECONOMIC GROWTH

India's recent economic growth has been closely intertwined with large-scale digitalisation driven by state-led digital public infrastructure (DPI). Artificial Intelligence (AI) has emerged as a critical layer within this digital ecosystem, shaping productivity, governance, service delivery, and global competitiveness. This article examines the role of AI in India's digitalisation process, situating it within the broader political economy of economic growth. It argues that AI functions less as an autonomous driver of growth and more as a force multiplier that enhances the effectiveness of digital infrastructure, state capacity, and market integration. However, the growth-enhancing potential of AI remains contingent on addressing structural constraints such as digital divides, skill gaps, and governance deficits.

India's economic trajectory in the twenty-first century has been marked by a strategic shift from purely factor-driven growth towards productivity-led expansion supported by digital technologies. Initiatives under the *Digital India* framework have created one of the world's largest digital ecosystems, encompassing biometric identity (Aadhaar), digital payments (UPI), tax administration (GSTN), and interoperable data frameworks. Artificial Intelligence (AI) has increasingly been embedded within this ecosystem, shaping how data is processed, decisions are made, and services are delivered.

Economic theory identifies productivity enhancement as a central channel through which technology contributes to long-term growth. In the Indian context, AI-driven analytics, automation, and predictive systems have begun to improve productivity across manufacturing, logistics, and services. AI-enabled predictive maintenance, quality control, and supply-chain optimisation allow firms to reduce wastage and downtime, enabling movement up the value chain ([Dundar et al., 2014](#)).

In manufacturing sectors aligned with Industry 4.0, AI complements digitalisation by integrating sensor data, machine learning, and real-time decision-making. This shift is particularly important for India as rising wages and global competition limit the sustainability of labour-intensive growth models.

A distinctive feature of India's digitalisation is the creation of state-backed digital public infrastructure. Platforms such as Aadhaar and UPI generate large volumes of standardised data, enabling AI-based applications in authentication, fraud detection, and service targeting. AI-driven systems improve the efficiency of welfare delivery by minimising leakages and exclusion errors, thereby enhancing fiscal efficiency and developmental outcomes (GOI, 2020). In financial services, AI-powered credit assessment models using digital transaction histories have expanded access to credit for MSMEs and informal-sector actors. This deepening of financial inclusion supports aggregate demand and investment, contributing to growth at the grassroots level ([World Economic Forum, 2024](#)).

Beyond markets, AI plays an increasingly important role in strengthening state capacity—an often-overlooked determinant of economic growth in developing economies. Predictive analytics is being applied in tax compliance, customs administration, and infrastructure planning to improve enforcement and resource allocation. Such applications reduce administrative inefficiencies and improve policy implementation, addressing long-standing governance bottlenecks ([Schiff, 2021](#)). In this sense, AI does not merely digitise existing processes but transforms governance by enabling data-driven decision-making. Improved state capacity, in turn, supports a stable macroeconomic environment conducive to sustained growth.

India's services-led growth model is undergoing transformation under the influence of AI. The traditional IT and BPO sectors, once driven by labour arbitrage, are increasingly shifting towards high-value AI services, data analytics, and platform-based solutions. This transition allows India to remain competitive in global markets while generating higher-skilled employment ([Jaiswal and Arun, 2021](#)). However, this shift also accentuates skill polarisation. While AI creates demand for advanced technical and cognitive skills, it threatens routine and low-skill employment. Without large-scale investment in skilling and reskilling, AI-driven growth risks becoming exclusionary, undermining its long-term sustainability.

AI's growth-enhancing potential is constrained by persistent digital divides across regions, genders, and income groups. Evidence from rural India demonstrates limited access to devices, connectivity, and digital literacy, particularly among women and marginalised communities ([ASER, 2021](#)). In such contexts, AI adoption risks reinforcing existing inequalities by disproportionately benefiting private institutions and urban elites ([Shalini and Tewari, 2020](#)). Thus, while AI can support inclusive growth—especially in agriculture, healthcare, and education—its distributive outcomes depend on complementary public investment and regulatory oversight.

AI governance remains a critical challenge. India's regulatory framework lacks sector-specific guidelines for AI deployment in areas such as education, welfare, and finance. Concerns regarding data privacy, algorithmic bias, and market concentration raise questions about the long-term developmental consequences of unregulated AI diffusion ([Global Information Society Watch, 2019](#)). At the strategic level, AI capability is increasingly linked to economic sovereignty. Dependence on foreign AI platforms risks data extraction and technological lock-in, prompting India to emphasise open, interoperable, and sovereign digital systems aligned with national growth objectives. Artificial Intelligence plays a significant but conditional role in India's digitalisation-led economic growth. It enhances productivity, strengthens state capacity, deepens financial inclusion, and supports India's transition towards a knowledge-based economy. However, AI is not a self-sustaining engine of growth. Its developmental impact depends on complementary investments in digital infrastructure, skills, and governance.

Table 1. India's Digitalisation, AI Enablers, and Economic Growth Linkages.

Indicator	Year(s)	Data	Growth / Economic Relevance	Source

Indicator	Year(s)	Data	Growth / Economic Relevance	Source
India GDP (current USD)	2000 → 2023	USD 468 bn → USD 3.6 tn	Shows long-run economic expansion coinciding with digital transition	World Bank
Digital Economy Share of GDP	2014 → 2023	~4.5% → ~11.7%	Digitalisation increasingly central to growth model	MeitY / McKinsey
Aadhaar Coverage	2010 → 2024	0 → 1.3+ billion IDs	Foundational data layer enabling AI-based governance	UIDAI
UPI Transactions (annual)	2017 → 2024	0.9 bn → 117 bn	AI-enabled fraud detection & credit analytics	RBI / NPCI
AI Adoption in Enterprises	2018 → 2023	~20% → ~55% firms	AI diffusion across productivity sectors	IBM Global AI Index
Contribution of Services to GDP	2000 → 2023	49% → 54%	AI-driven upgrading of IT–BPO and services	World Bank
MSMEs Using Digital Credit	2016 → 2023	~10% → ~45%	AI credit scoring boosts financial inclusion	WEF (2024)
Tax Buoyancy (GST era)	2017 → 2023	0.9 → 1.3	AI analytics improves compliance & revenue	GOI (GSTN)
Rural Internet Penetration	2014 → 2023	~18% → ~44%	Constraint on inclusive AI-led growth	ASER (2021)
Female Internet Usage (Rural)	2023	~33%	Digital divide limits AI benefits	ASER (2021)

The table demonstrates that India's economic growth acceleration coincides temporally with digital public infrastructure expansion, not AI in isolation. This empirically supports your claim that AI is *embedded within*—not independent of—digitalisation. Indicators such as UPI growth, tax buoyancy, and MSME credit expansion show that AI enhances the effectiveness of existing platforms, reinforcing your argument that AI amplifies institutional and market efficiency rather than directly generating growth. Data on digital divides (rural and gendered) empirically anchors the cautionary note that AI-led growth is structurally constrained by access, skills, and governance gaps, validating the conditional nature of AI's growth impact.

3 ARTIFICIAL INTELLIGENCE AND STRUCTURAL REFORM IN INDIA

India's recent economic growth has been accompanied by a gradual but significant process of structural reform aimed at improving productivity, state capacity, and market efficiency. Artificial Intelligence has emerged as an enabling technology within this reform trajectory, influencing governance, industrial organisation, labour markets, and service delivery. This

article examines whether and how AI contributes to India's structural reform agenda rather than merely acting as a technological add-on. It argues that AI functions as a reform multiplier by strengthening institutional effectiveness, reducing transaction costs, and facilitating formalisation, while also introducing new challenges related to inequality, employment displacement, and regulatory capacity.

Structural reform refers to changes in the institutional, regulatory, and productive foundations of an economy that enhance long-term growth potential rather than short-term cyclical performance. In India, such reforms have historically been slow, uneven, and constrained by administrative capacity, informality, and fragmented markets. Since the 2010s, however, India's reform strategy has increasingly relied on digitalisation as an institutional lever, with AI emerging as a key layer within this transformation.

Unlike earlier reform waves focused on deregulation and liberalisation, contemporary reforms in India increasingly emphasise governance efficiency, formalisation, and productivity enhancement, areas where AI has growing relevance. This article explores AI's role in India's structural reform process and assesses its implications for economic growth.

A central constraint on India's growth has been weak state capacity rather than absence of policy intent. AI contributes to structural reform by improving the effectiveness of governance through data-driven decision-making, predictive analytics, and automation. In tax administration, customs, and subsidy delivery, AI-enabled systems assist in risk profiling, fraud detection, and compliance monitoring, reducing discretion and leakages ([Schiff, 2021](#)). For instance, AI applications layered over digital tax platforms enhance revenue mobilisation without raising tax rates, thereby supporting fiscal sustainability—an essential precondition for long-term growth. This represents a qualitative shift in reform logic: instead of rewriting rules alone, AI enables better enforcement of existing rules, addressing a classic weakness of India's reform implementation.

India's economy has long been characterised by informality and fragmented markets, limiting productivity growth. AI contributes to structural reform by facilitating market integration through digital platforms and data interoperability. AI-powered credit assessment, logistics optimisation, and supply-chain analytics reduce information asymmetries and transaction costs, enabling small firms to participate in formal markets ([World Economic Forum, 2024](#)).

The formalisation of economic activity—accelerated by digital payments and AI-based compliance tools—expands the tax base and improves labour protections. While formalisation alone does not guarantee growth, it creates conditions for productivity-enhancing investment and scale economies, aligning with long-term structural reform objectives.

AI plays an increasingly important role in India's industrial restructuring, particularly in manufacturing and infrastructure. AI-enabled predictive maintenance, quality control, and energy management systems enhance capital efficiency and reduce operational losses. These technologies support India's transition from labour-intensive production to value-added, data-driven manufacturing, which is essential for sustaining growth amid demographic and global competitive pressures ([Dundar et al., 2014](#)). In this sense, AI complements broader reforms such as production-linked incentive (PLI) schemes by improving firm-level capabilities rather than merely providing fiscal incentives. Structural reform thus becomes embedded at the level of production processes.

From a structural perspective, AI introduces both opportunities and tensions within India's labour market. While AI-driven growth can generate high-skill employment in technology, analytics, and services, it also threatens routine and low-skill jobs, particularly in clerical, BPO, and manufacturing segments ([Jaiswal and Arun, 2021](#)). This creates a reform dilemma: AI enhances productivity and growth but risks undermining employment absorption in an economy with a large working-age population. Without parallel reforms in education, skilling, and labour market flexibility, AI-led structural change may exacerbate inequality and informalisation rather than resolve them.

Despite its reform potential, AI also introduces new structural risks. The absence of robust AI governance frameworks raises concerns regarding data privacy, algorithmic bias, and accountability. India's current regulatory architecture remains underdeveloped in addressing automated decision-making in welfare, finance, and employment ([Global Information Society Watch, 2019](#)). Moreover, market concentration in AI platforms risks creating new monopolies, potentially undermining competition—an outcome contrary to reform objectives. These challenges highlight that AI-driven reform is not self-correcting and requires strong institutional oversight.

Table 2. Artificial Intelligence as an Enabler of Structural Reform in India.

Structural Reform Dimension	Reform Indicator (Pre-AI → AI Phase)	Empirical Trend / Data	AI's Functional Role	Growth / Reform Implication	Source
State Capacity & Governance	Tax compliance & enforcement	GST tax buoyancy: ~0.9 (2017) → ~1.3 (2023)	AI-based risk profiling, anomaly detection, compliance analytics	Higher revenue without raising rates; improved fiscal sustainability	GOI (GSTN), Schiff (2021)
Formalisation of Economy	Digital payments share	Cashless transactions share <10% (2016) → >65% (2023)	AI fraud detection, transaction pattern analysis	Expansion of formal economic activity; lower transaction costs	RBI, NPCI
Market Integration	MSME access to formal credit	Digitally assessed MSME loans: ~10% (2016) → ~45% (2023)	AI credit scoring using transaction & GST data	Inclusion of small firms into formal markets; investment growth	World Economic Forum (2024)
Industrial Productivity	Manufacturing efficiency	Downtime reduction in AI-enabled plants: 15–25%	Predictive maintenance, quality control	Capital efficiency; shift to value-added manufacturing	Dundar et al. (2014)
Industrial Policy (PLI)	Firm-level capability upgrading	AI adoption among PLI-linked firms: ~50% (2023 est.)	Process optimisation, yield improvement	Structural upgrading beyond fiscal incentives	MeitY
Labour Market Structure	Skill composition of employment	High-skill tech jobs ↑; routine clerical jobs ↓	Automation, task reallocation	Productivity gains with employment polarisation risk	Jaiswal & Arun (2021)

Structural Reform Dimension	Reform Indicator (Pre-AI → AI Phase)	Empirical Trend / Data	AI's Functional Role	Growth / Reform Implication	Source
Agricultural Reform	Advisory & market access	AI-based advisory reach: ~25–30 million farmers	Weather, price, and crop analytics	Reduced uncertainty; productivity stabilisation	NITI Aayog
Human Capital Formation	Health & education access	AI pilots in telemedicine & ed-tech	Diagnostic support, personalised learning	Indirect growth via human capital	Shalini & Tewari (2020)
Competition & Market Structure	Platform concentration	Rising data-driven market power	Algorithmic scale advantages	Risk of monopolisation without regulation	GIS Watch (2019)
Economic Sovereignty	Dependence on foreign AI	High import dependence in core AI models	Push for indigenous/open AI systems	Alignment of reform with long-term autonomy	GOI (2020)

The table empirically demonstrates that AI does not initiate reforms, but enhances their effectiveness—especially in enforcement, compliance, and productivity—supporting your core theoretical claim. By linking AI to tax buoyancy, formalisation, MSME credit, and industrial efficiency, the table moves the discussion from abstract reform rhetoric to observable institutional change. Inclusion of labour polarisation, market concentration, and governance risks ensures analytical balance, strengthening credibility of the argument that AI enhance effectiveness.

4 ARTIFICIAL INTELLIGENCE AND INDIA'S BIG DATA ECOSYSTEM

India's economic growth in the past decade has been accompanied by the rapid expansion of a national big data ecosystem, driven by large-scale digitalisation of identity, finance, governance, and markets. Artificial Intelligence has emerged as a critical analytical layer within this ecosystem, transforming raw data into actionable economic and policy insights. This article examines the role of AI in shaping India's big data architecture and assesses its implications for productivity, governance, and long-term economic growth. It argues that AI functions as a growth-enabling intermediary—converting scale into value—while simultaneously raising challenges related to data governance, inequality, and market concentration.

Big data has become a foundational resource in contemporary economic systems, enabling real-time decision-making, efficiency gains, and innovation-driven growth. In India, the emergence of a large-scale big data ecosystem is closely tied to state-led digitalisation initiatives such as Aadhaar, Unified Payments Interface (UPI), Goods and Services Tax Network (GSTN), and sector-specific digital platforms in health, education, and agriculture. These systems generate unprecedented volumes of structured and semi-structured data. Artificial Intelligence plays a decisive role in this ecosystem by enabling the extraction of economic value from data. Without AI-driven analytics, machine learning, and predictive modelling, the scale of India's data would remain largely inert. This article analyses how AI interacts with India's big data ecosystem and evaluates its contribution to economic growth.

India's big data ecosystem is distinguished by its population-scale coverage, state-backed digital public infrastructure, and interoperability. Aadhaar alone covers over a billion individuals, while UPI processes billions of monthly transactions, creating continuous data flows on consumption, payments, and financial behaviour. The GSTN integrates firm-level transaction data across states, offering insights into supply chains and market integration (GOI, 2020). This ecosystem differs from purely market-driven data regimes by embedding public purpose within data generation. However, the economic value of such large datasets depends critically on AI-enabled processing and interpretation.

AI constitutes the analytical backbone of India's big data ecosystem. Machine learning algorithms enable pattern recognition, anomaly detection, and predictive analytics at a scale impossible through conventional statistical methods. In financial services, AI processes transaction-level data from digital payments to assess credit risk, detect fraud, and personalise financial products, thereby supporting financial deepening and MSME growth ([World Economic Forum, 2024](#)). From a growth perspective, AI-driven analytics reduce information asymmetries and transaction costs—two long-standing constraints on productivity in India's fragmented markets. This conversion of data into efficiency gains illustrates AI's indirect but substantial contribution to economic expansion.

AI's integration into big data systems has significant implications for governance and state capacity. Predictive analytics applied to tax data, welfare databases, and infrastructure systems enhances targeting accuracy, compliance, and policy effectiveness. For example, AI-enabled risk profiling in tax administration improves revenue mobilisation without increasing statutory tax rates, strengthening fiscal foundations for growth ([Schiff, 2021](#)). Moreover, AI-driven insights from big data enable evidence-based policymaking, allowing the state to anticipate economic shocks, regional disparities, and sectoral stress. This represents a qualitative shift from reactive governance to anticipatory regulation—an important structural support for sustained economic growth.

AI-driven big data analytics increasingly shape productivity across key sectors. In agriculture, AI combines satellite data, weather patterns, and market information to generate real-time advisory systems, addressing structural inefficiencies such as yield uncertainty and price volatility. In logistics and infrastructure, AI optimises routing, energy use, and asset maintenance, improving capital efficiency ([Dundar et al., 2014](#)).

In the services sector, particularly IT and platform-based businesses, India's comparative advantage increasingly lies in data-intensive and AI-enabled services rather than labour

arbitrage alone. This transition supports higher value-added growth and global competitiveness.

While AI enhances the economic value of big data, it also reshapes market structures. Firms with superior access to data and advanced AI capabilities enjoy increasing returns to scale, raising concerns about market concentration and entry barriers. In India, this dynamic risks creating data monopolies that could undermine competition and innovation—counteracting growth-enhancing reform objectives ([Global Information Society Watch, 2019](#)). Public digital infrastructure partially mitigates this risk by democratising access to foundational data. However, the analytical advantages conferred by proprietary AI systems remain unevenly distributed.

AI-driven big data ecosystems can either support inclusive growth or reinforce existing inequalities. While digital platforms generate data across socio-economic groups, the benefits of AI analytics disproportionately accrue to entities with technological and analytical capacity. Moreover, gaps in digital access and data literacy limit the participation of marginalised communities, particularly in rural areas and among women ([ASER, 2021](#)). Thus, the growth impact of AI-enabled big data depends on complementary investments in digital infrastructure, skills, and open-access analytical tools.

The expansion of AI-driven big data systems raises critical governance challenges. India's regulatory framework for data protection, algorithmic accountability, and consent remains evolving. Weak governance can undermine trust, distort markets, and generate economic inefficiencies through misuse of data and biased decision-making ([Global Information Society Watch, 2019](#)). From a growth perspective, robust data governance is not merely an ethical concern but an economic necessity, ensuring sustainable and predictable returns from data-driven innovation.

At a strategic level, AI capabilities determine whether data-driven growth accrues domestically or is captured by global platforms. Dependence on foreign AI technologies risks value extraction and technological lock-in. India's emphasis on open, interoperable digital systems reflects an effort to align big data utilisation with national economic sovereignty and long-term growth objectives (GOI, 2020). Artificial Intelligence plays a pivotal but conditional role in India's big data ecosystem. It transforms scale into economic value by enabling productivity gains, strengthening governance, and supporting market integration. However, AI is not an automatic engine of growth. Its contribution depends on inclusive data access, robust governance frameworks, and strategic public investment.

Table 3. Artificial Intelligence within india's big data ecosystem and economic growth.

Big Data Dimension	Core Data Infrastructure	Empirical Indicator / Trend	Role of AI	Growth / Political Economy Implication	Source

Big Data Dimension	Core Data Infrastructure	Empirical Indicator / Trend	Role of AI	Growth / Political Economy Implication	Source
Population-Scale Data Generation	Aadhaar	Coverage: 0 (2009) → 1.3+ bn individuals (2024)	Biometric matching, anomaly detection	Enables scalable service delivery and trust in markets	UIDAI; GOI
Transaction Data & Payments	UPI	Annual transactions: <1 bn (2017) → 117 bn+ (2024)	Fraud detection, transaction pattern analysis	Reduced transaction costs; market deepening	RBI; NPCI
Firm-Level Market Data	GSTN	Nationwide firm-level transaction visibility (post-2017)	Risk profiling, compliance analytics	Supply-chain integration; revenue efficiency	GOI (GSTN)
Financial Deepening	Digital payments + GST + bank data	Digitally assessed MSME credit: ~10% (2016) → ~45% (2023)	AI credit scoring & cash-flow analytics	MSME formalisation; investment growth	World Economic Forum (2024)
Governance & State Capacity	Welfare, tax, customs databases	Improved targeting; reduced leakages	Predictive analytics; anomaly detection	Shift from reactive to anticipatory governance	Schiff (2021)
Agricultural Data Ecosystem	Satellite, weather, mandi prices	AI advisories reach ~25–30 million farmers	Yield prediction; price forecasting	Risk reduction; productivity stabilisation	NITI Aayog
Infrastructure & Logistics	Transport & energy data	AI optimisation reduces downtime 15–25%	Route, load, and energy optimisation	Capital efficiency; lower logistics costs	Dundar et al. (2014)

Big Data Dimension	Core Data Infrastructure	Empirical Indicator / Trend	Role of AI	Growth / Political Economy Implication	Source
Services & Digital Platforms	IT, e-commerce, fintech data	Shift from BPO to data-driven services	ML-based analytics & platforms	Higher value-added services growth	Jaiswal & Arun (2021)
Market Structure	Platform data concentration	Increasing returns to scale	Algorithmic advantage from data accumulation	Risk of monopolisation; entry barriers	GIS Watch (2019)
Inclusion & Data Divide	Rural & gendered access	Rural internet ~44%; rural women ~33%	AI benefits unevenly distributed	Risk of inequality-reinforcing growth	ASER (2021)
Data Governance	Consent, privacy, accountability	Fragmented regulatory framework	Algorithmic opacity	Trust deficit; economic inefficiencies	GIS Watch (2019)
Economic Sovereignty	Core AI models & platforms	High dependence on foreign AI stacks	Strategic push for open/indigenous AI	Retention of domestic value creation	GOI (2020)

The table shows that India's big data abundance alone does not generate growth; AI is the mechanism that converts identity, transaction, and firm-level data into productivity, inclusion, and governance gains. By including market concentration, data governance, and sovereignty, the table avoids technological determinism and embeds AI within institutional and power structures. Rows on digital divides and governance risks empirically support your claim that AI-enabled big data can either support inclusive growth or reinforce inequality, depending on policy choices.

5 THE ROLE OF ARTIFICIAL INTELLIGENCE IN INDIA'S REMARKABLE ECONOMIC STRIDE IN RECENT DECADES

India's sustained economic expansion over the past decades reflects a complex interaction of demographic advantage, policy reform, global integration, and technological change. In recent years, Artificial Intelligence has emerged as an important enabling technology within this growth narrative. Rather than acting as a standalone driver of economic expansion, AI has functioned as a force multiplier—enhancing productivity, strengthening state capacity,

accelerating market integration, and upgrading India's position in global value chains. This article analyses the role of AI in India's economic stride, situating it within the country's digital transformation and political economy of growth, while highlighting the structural constraints that shape its long-term impact.

India's economic trajectory since the late twentieth century has been characterised by gradual structural transformation, punctuated by policy reforms and technological shifts. Liberalisation in the 1990s integrated India into global markets, while the 2000s witnessed services-led growth driven by information technology. In the past decade, this growth pattern has been reinforced by large-scale digitalisation and the emergence of Artificial Intelligence as a general-purpose technology. AI's relevance to India's economic stride lies not merely in automation or innovation but in its capacity to convert scale into efficiency—a critical requirement in a large, diverse, and institutionally complex economy. This article examines how AI contributes to India's growth momentum by enhancing productivity, governance, and competitiveness, while also acknowledging the limits and risks associated with uneven technological diffusion.

Productivity growth is central to sustained economic expansion, particularly for an economy transitioning from factor-driven to efficiency-driven growth. AI-enabled analytics, automation, and optimisation tools have begun to improve productivity across manufacturing, logistics, finance, and services. In Indian manufacturing, AI supports predictive maintenance, quality control, and energy efficiency, helping firms move up the value chain amid rising labour costs and global competition ([Dundar *et al.*, 2014](#)). In services, AI augments India's traditional strengths in IT by shifting the sector towards higher-value activities such as data analytics, machine learning services, and platform-based solutions. This transition allows India to retain global competitiveness while generating productivity gains that contribute directly to GDP growth.

A distinctive feature of India's recent economic stride is the creation of population-scale digital public infrastructure (DPI). Platforms such as Aadhaar, UPI, and GSTN generate vast datasets on identity, payments, and transactions. AI acts as the analytical layer that transforms this data into economic value through fraud detection, credit scoring, logistics optimisation, and policy targeting (GOI, 2020). The growth impact of AI-enabled DPI is indirect but significant. By reducing transaction costs, improving trust, and enabling market integration, AI strengthens the institutional foundations of growth rather than merely stimulating short-term output.

Weak implementation capacity has historically constrained India's growth potential. AI contributes to economic performance by improving governance efficiency through predictive analytics, automated compliance monitoring, and evidence-based policymaking. In tax administration and subsidy delivery, AI-assisted systems reduce leakages and discretion, enhancing fiscal efficiency without increasing distortionary taxation ([Schiff, 2021](#)). Improved state capacity supports macroeconomic stability and investor confidence—critical conditions for sustained economic growth in a large emerging economy.

AI has played an important role in financial deepening, a key driver of India's recent growth. AI-driven credit assessment models using digital transaction data have expanded access to finance for MSMEs and informal-sector actors, supporting entrepreneurship and consumption-led growth ([World Economic Forum, 2024](#)). By reducing information

asymmetries and expanding formal credit access, AI strengthens domestic demand and investment—two pillars of India's growth momentum in recent decades.

While AI enhances productivity, it also introduces trade-offs in labour-intensive sectors. Automation threatens routine employment in clerical, BPO, and low-skill manufacturing segments, raising concerns about jobless growth. At the same time, AI creates demand for high-skill labour in data science, engineering, and analytics ([Jaiswal and Arun, 2021](#)). India's ability to sustain its economic stride therefore depends on complementary investments in education, skilling, and reskilling, without which AI-driven growth risks becoming socially and politically unsustainable.

AI's contribution to growth is uneven across regions and social groups. Urban centres with advanced digital infrastructure benefit disproportionately from AI-led innovation, while rural and marginalised communities face barriers related to connectivity and digital literacy ([ASER, 2021](#)). This uneven diffusion risks widening regional and social inequalities, potentially undermining aggregate growth through demand constraints. Inclusive AI deployment—particularly in agriculture, healthcare, and education—is therefore essential for sustaining India's growth trajectory.

AI has become a strategic economic asset in global competition. India's growing AI capabilities enhance its position in global value chains and reduce dependence on foreign technology platforms. However, reliance on imported AI tools risks data extraction and value capture by global firms. India's emphasis on open, interoperable digital systems reflects an attempt to align AI adoption with long-term economic sovereignty ([Global Information Society Watch, 2019](#)).

Artificial Intelligence has played a supporting yet increasingly significant role in India's remarkable economic stride in recent decades. By enhancing productivity, strengthening governance, deepening financial markets, and upgrading services, AI reinforces the structural foundations of economic growth. However, AI is not a substitute for sound policy, institutional reform, or human capital development.

Table 4. Artificial intelligence and india's economic stride in recent decades.

Growth Dimension	Structural Context (Pre-AI)	Empirical Trend (AI Phase)	Role of AI	Growth Implication	Source
Overall Economic Expansion	GDP growth driven by labour & capital accumulation	GDP: ~USD 468 bn (2000) → USD 3.6 tn+ (2023)	Efficiency gains across sectors	Transition from factor-to productivity-led growth	World Bank

Growth Dimension	Structural Context (Pre-AI)	Empirical Trend (AI Phase)	Role of AI	Growth Implication	Source
Productivity Enhancement	Low total factor productivity (TFP) growth	Manufacturing downtime ↓ 15–25% in AI-enabled plants	Predictive maintenance; quality analytics	Higher output per unit of capital	Dundar <i>et al.</i> (2014)
Digital Public Infrastructure (DPI)	Fragmented identity & payments	Aadhaar (1.3 bn IDs); UPI (117 bn txns, 2024)	Fraud detection; credit & logistics analytics	Lower transaction costs; market integration	GOI (2020); RBI
State Capacity & Governance	Weak enforcement, high leakages	GST buoyancy: ~0.9 (2017) → ~1.3 (2023)	Risk profiling; compliance automation	Fiscal efficiency without higher taxes	Schiff (2021)
Financial Deepening	Limited MSME credit access	Digitally assessed MSME loans: ~10% (2016) → ~45% (2023)	AI credit scoring using transaction data	Investment and consumption-led growth	WEF (2024)
Services-Led Growth	IT–BPO driven by labour arbitrage	Shift towards AI, analytics & platforms	ML-based service upgrading	Higher value-added exports	Jaiswal & Arun (2021)
Industrial Upgrading	Labour-intensive manufacturing	AI adoption in Industry 4.0 clusters	Process optimisation; energy efficiency	Movement up global value chains	MeitY
Labour Market Structure	High absorption of routine labour	High-skill jobs ↑; routine jobs ↓	Task automation & reallocation	Productivity gains with employment risks	Jaiswal & Arun (2021)
Regional & Social Inclusion	Urban-centric growth	Rural internet ~44%; rural women ~33%	Uneven AI diffusion	Risk of inequality-driven growth limits	ASER (2021)

Growth Dimension	Structural Context (Pre-AI)	Empirical Trend (AI Phase)	Role of AI	Growth Implication	Source
Global Competitiveness	Dependence on services outsourcing	Rising AI service exports	Data-intensive competitiveness	Strategic upgrading of growth model	OECD
Economic Sovereignty	Reliance on foreign tech platforms	Push for open, interoperable AI systems	Indigenous AI & DPI alignment	Retention of domestic value creation	GIS Watch (2019)

The table explicitly links AI adoption to India's decades-long growth transition, reinforcing your argument that AI is embedded in structural change rather than episodic innovation. Instead of overstating AI's role, the table shows how AI strengthens existing growth drivers—productivity, DPI, state capacity, and services upgrading.

6 IS AI A BOON OR A BANE?

India's remarkable economic expansion in recent decades has been shaped by a confluence of demographic dynamics, policy reform, global integration, and technological change. Since the late twentieth century, India's growth trajectory has moved gradually from factor-driven expansion toward productivity-led growth, supported increasingly by digitalisation and institutional reform. Within this transformation, Artificial Intelligence has emerged not as an autonomous driver of growth, but as an enabling technology that amplifies the effectiveness of digital infrastructure, structural reforms, and data-driven governance.

Digitalisation has become a central pillar of India's development strategy, particularly since the 2010s. State-led digital public infrastructure—including Aadhaar, the Unified Payments Interface, and the Goods and Services Tax Network—has created one of the world's largest interoperable digital ecosystems. These platforms have reduced transaction costs, improved market integration, and strengthened trust in economic exchanges ([Government of India, 2020](#)). AI functions as a critical analytical layer within this ecosystem, transforming large volumes of administrative and transactional data into actionable insights through machine learning, predictive analytics, and automation. In this sense, AI converts digital scale into economic efficiency rather than independently generating growth.

Structural reform in India has historically been constrained by weak implementation capacity, high informality, and fragmented markets. Recent reform strategies have increasingly relied on digitalisation to address these constraints. AI contributes to this process by improving governance effectiveness rather than altering formal rules alone. In tax administration, welfare delivery, and customs enforcement, AI-enabled risk profiling and anomaly detection enhance compliance and reduce leakages, strengthening fiscal capacity

without raising distortionary tax rates ([Schiff, 2021](#)). This represents a qualitative shift in India's reform logic, where enforcement and institutional effectiveness are improved through data-driven tools rather than solely through legislative change.

India's expanding big data ecosystem provides the material foundation for AI-enabled reform and growth. Population-scale data generated through digital identity systems, payment platforms, and firm-level tax records create unprecedented opportunities for economic coordination and policy targeting. AI enables pattern recognition and predictive modelling at a scale that conventional statistical methods cannot achieve. In finance, AI-based credit assessment using digital transaction histories has expanded access to formal credit for micro, small, and medium enterprises (MSMEs), supporting entrepreneurship, investment, and consumption-led growth ([World Economic Forum, 2024](#)). By reducing information asymmetries, AI strengthens financial deepening—an important driver of India's growth momentum.

At the sectoral level, AI contributes to productivity enhancement across manufacturing, logistics, agriculture, and services. In manufacturing, AI-enabled predictive maintenance and quality control improve capital efficiency and support movement up the value chain, which is increasingly necessary as labour-cost advantages erode ([Dundar *et al.*, 2014](#)). In services, AI is reshaping India's traditional IT–BPO model by shifting it toward higher-value activities such as data analytics and platform-based services, thereby sustaining global competitiveness ([Jaiswal and Arun, 2021](#)). These productivity gains reinforce long-term growth rather than producing short-term output spikes.

However, AI's contribution to India's economic stride remains uneven and conditional. Digital divides across regions, genders, and income groups limit the inclusive diffusion of AI-enabled benefits ([ASER, 2021](#)). Moreover, AI-driven automation poses risks of labour displacement in routine and low-skill occupations, raising concerns about jobless growth in a labour-abundant economy. Governance challenges—including data privacy, algorithmic bias, and market concentration—further complicate AI's developmental impact ([Global Information Society Watch, 2019](#)). Without complementary investments in education, reskilling, and regulatory capacity, AI risks reinforcing existing inequalities rather than alleviating them.

Artificial Intelligence has played a meaningful but supportive role in India's remarkable economic stride in recent decades. It operates primarily as a force multiplier—enhancing the productivity, governance, and market-integration effects of digitalisation, structural reform, and big data ecosystems. AI is not a substitute for sound policy, institutional reform, or human capital development. Its long-term contribution to India's economic growth depends on how effectively it is embedded within inclusive digital infrastructure, robust governance frameworks, and a broader strategy of sustainable development.

7 CONCLUSION

The evidence suggests that AI has contributed indirectly but meaningfully to India's economic stride by strengthening governance, deepening financial inclusion, upgrading services, and facilitating market integration. Population-scale platforms such as Aadhaar, UPI, and GSTN provide the foundational data architecture, while AI enables these systems to generate

economic value through analytics, automation, and predictive decision-making. In this sense, AI reinforces the institutional foundations of growth rather than substituting for policy reform, human capital development, or macroeconomic stability.

At the same time, the analysis highlights clear limits and risks. AI-driven growth remains uneven across regions and social groups, reflecting persistent digital divides, skill gaps, and infrastructural disparities. Labour market disruptions, concerns over data governance, algorithmic bias, and market concentration pose significant challenges to inclusive and sustainable growth. Without complementary investments in education, reskilling, regulatory capacity, and open digital governance, AI risks reinforcing existing inequalities rather than alleviating them.

In conclusion, Artificial Intelligence has played a supportive yet increasingly significant role in India's remarkable economic progress in recent decades. Its developmental impact is conditional, mediated by institutions, governance frameworks, and social policy choices. AI should therefore be understood not as a substitute for structural reform, but as an enabling layer that can strengthen reform outcomes when embedded within a coherent strategy of inclusive growth, state capacity-building, and economic sovereignty.

REFERENCES

- ASER (2021) *Did Technology Enable Inclusive Education in Rural India?* [↗](#)
- Dundar, H. et al.. (2014) *Student Learning in South Asia: Challenges, Opportunities, and Policy Priorities*. World Bank. [↗](#)
- Global Information Society Watch (2019) *Artificial Intelligence: Human Rights and Social Justice*. [↗](#)
- Government of India (2020) *National Education Policy 2020*. [↗](#)
- Government of India (2020). *India's Digital Public Infrastructure and Data Empowerment*.
- Jaiswal, A. and Arun, C.J. (2021) 'Potential of Artificial Intelligence for Transformation of the Education System in India,' *IJEDICT*, 17(1), pp. 142–158. [↗](#)
- Schiff, D. (2021) 'Education for AI, not AI for education,' *International Journal of Artificial Intelligence in Education*. [↗](#)
- Shalini and Tewari, A. (2020) 'Sustainable education in india through artificial intelligence,' *WebSci Companion*. [↗](#)
- World Economic Forum (2024) *Shaping the Future of Learning: The Role of AI in Education 4.0*.
- World Economic Forum (2024) *The Economic Value of Big Data and AI*.
- World Economic Forum (2024) *The Role of AI in Economic Transformation*.
- World Economic Forum (2024). *AI, Big Data, and Financial Inclusion in Emerging Markets*. [↗](#)

Algorithmic Minds in the Classroom: AI, Identity, and the Future of Critical Thinking in Education

Garima Aggarwal and Ashish Ranjan

University of Delhi

DOI: [10.1201/9781042018116-18](https://doi.org/10.1201/9781042018116-18)

ABSTRACT: Artificial intelligence and its educational tools pose significant questions across the diverse landscape of education, specifically in countries like India. The paper attempts a ‘critical conceptual inquiry’ within an ‘interpretivist approach’ with a comprehensive overview of around 50 studies in the relevant areas, including pertinent policy documents from the apex bodies of India. It investigates how different AI platforms, through data-driven feedback, are generating an ‘unreflective mind’. It examines how this further impacts the ‘critical thinking skills’ and ‘epistemic selves’ of the students. The paper introduces the concept of ‘AI-algorithmic minds’, drawing upon psychological frames such as Vygotsky's socio-cultural perspective (1978), and Bandura's theory of self-efficacy (1997), sociological frames such as Foucault's theory of disciplinary power (1977), and critical theory foundations such as Paulo Freire's critical pedagogy (1970). The concept is further used to assess the learning patterns generated through the mediation of AI. The paper also attempts to understand how government policy initiatives (e.g., DIKSHA, PM eVidya, NDEAR, UDISE+) and apex bodies (NITI AYOOG, CBSE) adopt these global changes. The analysis reveals the proliferation of speed and efficiency at the cost of reflective learning by AI-driven technologies. This results in the marginalisation of local/ regional knowledge and exacerbates

digital inequalities. The paper further proposes the PACE+ framework (Participation, awareness, contextualisation, engagement, empathy, and slowness) to envisage a much-needed human intervention in the development of AI tools that supports a culturally sensitive, emotionally intelligent, pedagogically grounded, and ethically responsive intervention that is shaped by local values and educational justice.

Keywords: Artificial intelligence in education, socio-cultural theories, identity formation, critical thinking, and metacognition

1 INTRODUCTION

There is a rapid transformation of education through Artificial Intelligence across the globe. It impacts how education is delivered, experienced, comprehended, and assessed. Several models of AI practice in education are overtly relied upon, including AI technologies for personalized and automated learning, performance tracking, and smart tutoring. Many countries have launched AI-driven platforms, such as Squirrel AI in China and Knewton in the United States. These platforms tailor educational content to students based on their behavioral and performance needs.

Historically, the education system has been grappling to attain non-biased and systematic grading. These days, with the help of AI, tools like Gradescope are developed, which are considered efficient in streamlining these issues. AI is employed not just in transaction and assessment of education, rather it is actively used in providing conversational tutoring to the students using natural language interfaces such as the Google Socratic app. In higher education platforms such as Coursera and EdX, AI is used to recognise the pattern of the students' inclinations towards different courses, which is further used for recommendations. This supports students' retention (Lukin *et al.*, 2016).

The progression of education towards AI is rapid and huge. It is impractical to ignore the impact of the AI-driven education system at present. Yet, there are critical questions that have been left unanswered. Some of the relevant questions are: How do these AI tools influence the

way students see themselves as learners? What kind of cognitive habits are being encouraged when students receive instant feedback and answers generated by the way of AI algorithm? Will it be unintentionally creating environments where students are more engaged in speed instead of asking critical questions? What theoretical frames can be employed to understand the role of AI-intervention in Indian classrooms? How do the Indian government policy initiatives (e.g., DIKSHA, PM eVidya, NDEAR, UDISE+) and apex bodies (NITI AYOg, CBSE) adopt and understand these global changes? What should be the future direction to comprehend the role of AI, and what kind of interventions are needed?

This paper attempts to answer these questions and introduces the concept of ‘AI Algorithmic Minds’- a term which is coined in this paper in the context of Artificial Intelligence to describe students' learning/thinking that is shaped by data-driven patterns, recommendations, and scoring made by Artificial Intelligence.

2 INDIAN CONTEXT AI TECHNOLOGY

Artificial intelligence is increasingly being observed as a remarkable tool in education. Globally, researchers have talked about the usage of AI to increase personalisation of the learning experience, reduce the administrative load on the teachers, and offer continuous feedback to the students ([Luckin et al., 2016](#); [Homes et al., 2019](#)). According to bibliometric analysis, research in the field has increased at an estimated annual rate of 47.9% ([Zawacki-Richter et al., 2019](#)). Only within the US and China, the publication in the field sums up to 40% of all the publications ([Xia et al., 2022](#); [Zawacki-Richter et al., 2019](#)). Platforms like BYJU's, Toppr, and Vedantu are already using the AI algorithms to analyse students' responses and learning pace to provide personalized educational journeys. There is a kind of confidence that AI has generated for itself as an educational tool with the help of its fast-paced pattern-finding abilities and quick decision-making based on these patterns (Williamson and [Knox, 2020](#)). Most research in AI-enabled education systems has employed methods such as experimental designs and bibliometric overviews ([Chen et](#)

[al., 2020](#); [Zawacki-Richter et al., 2019](#)). Keywords landscape highlights terms such as personalized learning, adaptive systems, critical thinking, identity, agency, AI ethics, socio-cultural context, inclusion, and digital divide.

In the Indian context, Significant moves have been taken to integrate AI into the classrooms by many well-known schools (ET [Edge Insights., 2024](#); [IndiaAI, 2024](#); [India Today., 2025](#)). The shift towards digital platforms has accelerated even faster since the COVID-19 pandemic. The acceleration can be witnessed even under the governmental decisions for education, such as the declaration of the National program DIKSHA (The national digital infrastructure for teachers offers thousands of AI-supported resources for school education and professional development) and PM eVidya (launched in 2020, facilitates multi-platform digital learning, including the use of AI to analyse learning patterns and improve content delivery). The Central Board of Secondary Education (CBSE) has introduced AI as an elective subject in many schools. The board has kept a target of training over 20,000 teachers to handle AI-related content and pedagogy (CBSE Annual Report, 2022).

[INITI Aayog, in a paper titled “National Strategy for Artificial Intelligence” \(2018\)](#), emphasized that there is huge transformative potential in the education sector for AI in India. The paper suggested that AI could help in personalizing the educational experience, which can further impact the dropout ratio. However, data from the Unified Information System for Education Plus (UDISE+ is an Indian government data collection system on district-wise schools under the Department of School Education and Literacy. It is designed to collect real-time data to support planning, monitoring, and decision-making) (UDISE+, 2022) suggested that, despite movements like Digital India, there is still a wide gap in the digital infrastructure between the public schools and government schools in India. Only about 28 percent of schools in India have internet facilities as of 2021-22. India is a diverse country, where imparting educational experiences has always been rooted in community learning, storytelling, and dialogic methods. An uncritical adoption of AI in the classroom can put various communities at the periphery.

3 POLICY INITIATIVES IN THE INDIAN CONTEXT

DIKSHA has adopted AI integration to flag learning errors. This is further used to provide feedback and tailored content that targets the specific learning problem. The integration has been largely completed for mathematics and science lessons at the secondary stage, adopting a focused approach. It uses Personalised Adaptive Learning (PAL) to detect errors and modify the content accordingly. The platform also supports contributions in regional languages and has content created by local teachers. Hence, the standardization level is still moderate. However, it is worth noting that generating content in the regional language is not sufficient until locals are involved in it, allowing for connections with tribal and other minority languages. This leads to limited inclusivity. Further, the teacher-generated content is going through robust procedures, and delays in its approval are creating a gap cycle ([Banerjee and Biswas, 2025](#)).

PMeVidya unifies digital, TV (Swayam Prabha), radio, and other online platforms for a nationwide reach. The AI integration here has been done to derive content instead of using it to get personalised content, hence the standardisation is high. It also provides content in 30 Indian languages; however, the content is limited in other languages than Hindi and English ([Jain, 2025](#)).

NDEAR is not a content provider but is a framework that states can use to develop AI-powered solutions. It enables data sharing and supports community development. Limitations here also impede NDEAR's implementation. While the standards are provided for regional language support and integration, the real work in the field lags due to various practical restrictions such as script encoding, font support, dialectical variations, and limited regional content support (National Digital Education Architecture, 2022).

The implementation of these initiatives has a few difficulties, such as the infrastructural and digital divide that it creates. Many of the schools do not have adequate hardware and consistent electricity to implement these initiatives. That severely restricts the effectiveness of these programs and creates a digital divide between haves and have-nots. That further deepens the urban-rural gap ([Siddiq et al., 2024](#)). Another gap that can be observed

in the implementation of these programs is the lack of digital literacy among teachers in an appropriate manner. Teachers require substantial support in such integrations, and there is a notable lack of teacher education in this area ([Mehta, 2024](#)). Though the policy initiatives of the Indian government have been launched in various regional languages, maintaining a continuous generation of contextually appropriate content becomes a challenge with over-reliance on AI integration. Over a period of time, it also develops in lack of emotional connection and social skills development ([Mehta, 2024](#)). Also, the collection and management of data related to students through AI is a serious safety concern ([Hatkar, 2025](#)).

4 COGNITIVE OFFLOADING

Risko and Gilbert (2016) coined the term ‘cognitive offloading’ to illustrate how over-reliance upon digital tools can diminish internal cognitive processing. The shortcut to think that is supported by openly available AI tools to the learner can result in a non-questioning and conformist attitude. Such a learner chooses suggested answers and follows the recommended sources of knowledge uncritically.

One can wire the AI system to give a critical point of view as a result in their answer scripts. However, critical thinking is not just about getting the answers. It is about developing a mindset that can ask the right questions and think contextually. It requires curiosity and scepticism at the same time. These qualities are built by the education system over a period of time. They are not built by getting ready-made answers, even if they are critical in their responses ([Facione, 2015](#)).

5 MANIFESTATION IN INDIAN CLASSROOMS

Students these days are increasingly relying on AI-driven platforms for creating assignments and research work. This is generating a culture of ‘copy-paste’ that doesn’t require thinking and analytical skills on the part of the students. This reliance diminishes students’ capabilities to engage with

the content/ text meaningfully and be able to develop ‘ownership’ of learning. This further results in weaker brain engagement and reduced long-term recall power (Khan *et al.*, 2025). When asked to defend their work with ownership, they are unable to explain themselves with a ‘voice’.

Learning is a process that requires independence in thinking and a struggle that students go through to generate lifelong understandings and a zeal for further learning. AI-driven technologies create a shortcut for the students that defeats the very joy of learning as well ([Khan, Qamar, Ansari, and Yasmeen, 2025](#)).

6 UNDEREXPLORED RESEARCH LENSES

Despite the huge corpus of research that now exists on AI-driven education, limited work exists on the socio-cultural dimensions of AI tools. Ethical equity-based concerns, as well as a cognitive and critical perspective to understand what are the limitations of minds that are created in an AI-proliferated ecosystem, are addressed in fewer publications ([Xia *et al.*, 2022](#)). Artificial intelligence lacks context-based thinking and sensitivity that a human teacher can produce during pedagogical dialogues. Several issues, like student motivation and well-being, take a positive course after they engage with the real-life experiences of their mentors ([Yeager and Dweck, 2012](#); [Facione, 2015](#)). This is particularly significant in a country like India, where learning styles are highly diverse and are intertwined with regional, ethnic, and language-based differences ([UNESCO, 2023](#); NCERT, 2021).

The goal of this paper is fourfold. To introduce the concept of ‘AI Algorithmic Minds’ with the help of ‘theoretical frames’ that help in the investigation of cognitive shifts taking place in the classrooms. To understand the identity formation of the students in an epistemic context impacted by AI-driven technologies. To examine the Indian policy milieu and contextualise the AI-driven technological impact on its horizon. To outline the PACE+, a pedagogically grounded, emotionally intelligent, and culturally sensitive framework that envisages the intervention of humans for a more responsive AI technological platform for the learners. By grounding

our work in Indian realities and global research, we aim to contribute to a more reflective, human-centred dialogue around AI in education.

7 METHODOLOGY

The present paper adopts a ‘critical conceptual inquiry’ approach, guided by an ‘interpretivist framework’ (The interpretivist framework is rooted in the understanding that reality is socially constructed and is context-dependent. Here, cultural standpoint attains significance, and knowledge is not seen as neutral. For more details, please go through- Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin and Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 189–213). Sage Publications) to explore how Artificial Intelligence has the potential to shape the cognition and epistemic selves of the students. It further enquires how AI integrates with the socio-cultural component that significantly impacts education. ‘Critical conceptual inquiry’ as a qualitative approach refers to a rigorous examination of key ideas, relevant theoretical frameworks, and the assumptions behind the shaping of these ideas. This method presumes that any shift in society is not neutral; it is embedded in power structures and ideological assumptions ([Kincheloe, 2005](#)). This approach is suitable when the goal of the research is to introduce a new conceptual framework to offer reflective scholarship on real-life issues (Rocco and Plakotnik, 2009). Rather than collecting new empirical data, this approach synthesizes the existing scholarship but with a critical lens of enquiry, which was missing in the earlier conceptualization.

The paper presents an in-depth analysis of government policy data and classroom-level research, along with an inter-disciplinary framework to theorize the emerging phenomenon of ‘algorithmic minds’ in Artificial Intelligence. It asks the overlooked questions and critically conceptualises the far-reaching impacts of Artificial Intelligence and its pedagogical integration in education.

Research questions

The present inquiry is guided by three central questions:

- How do AI tools impact the learner's self-concept and critical thinking skills?
- Have AI-driven technologies been able to integrate the socio-cultural knowledge that is inevitably embedded in educational experiences?
- How educators can further utilize AI technology while retaining the pedagogical depth and cultural relevance of education.

8 THEMATIC SCOPE

The literature reviewed for this paper was selected through ‘purposive sampling’ (Purposive sampling” is a non-probability sampling technique used in qualitative research. This technique is used when the goal of the research is to gain an in-depth understanding of the experiences of the participants within a contextual framework. It selects participants who suit the goals, characteristics, and knowledge relevant to the research. For more details, please go through- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications) from scholarly databases such as JSTOR and ResearchGate. The theoretical anchors for the research were developed after a thorough literature review, and research questions were conceptualised. The theoretical framework of the present research is drawn from the socio-cultural learning perspective ([Vygotksky, 1978](#); [Rogoff, 2003](#)), epistemic identity and self-concept ([Yeager and Dweck, 2012](#)), critical pedagogy and social justice ([Smith, 1999](#); [Fricker, 2007](#)), disciplinary power ([Foucault, 1977](#)), and theories from cognitive psychology ([Flavell, 1979](#); Risko and Gilbert, 2016).

To contextualise the research within India, governmental data were drawn from the Ministry of Education UDISE+ (2021-22) report, CBSE Annual report (2022), and NITI Aayog's documentation on ‘National strategy for Artificial Intelligence’ (2018). This helps to contextualize the research in the socio-cultural reality the India.

Literature review was conducted from the year 2000 to 2024, with a specific focus on post-2015 literature. The themes that emerged during the literature review were AI-enabled learning and its pedagogical assumptions; development of epistemic identities and self-perception in students; socio-

cultural learning theories and their intersection with AI-enabled learning; power structures, surveillance, and marginalisation of the disadvantaged groups.

Table 1 1. Systematic review with stated metrics.

Lens/Topic	No of Papers (Approx.)	References
AI in Education: Reviews, Policy, and Trends	11	Chen et al. (2020) ; Holmes et al. (2019) ; Luckin et al. (2016) ; Xia et al. (2022) ; Williamson and Eynon (2020) ; Zawacki-Richter et al. (2019) ; Ministry of Education (2020, 2021, 2022) ; NITI Aayog (2018); UNESCO (2021, 2022, 2023) ; IndiaAI (2024) ; ET Edge Insights (2024)
Personalized Learning / Adaptive learning	7	Bardia and Agrawal (2025) ; BestColleges (2025) ; SSRVM (2025) ; Graesser et al. (2014) ; India Today (2025) ; IJARSCT (2024) ; Sanghvi (2024)
Critical Thinking	5	Facione (2015) ; Dewey (1933) ; Flavell (1979) ; Graesser et al. (2014) ; Kellner and Share (2007)
Identity, agency, self-concept	3	Smith (1999) ; Freire (1970) ; Yeager and Dweck (2012)
Socio-cultural and non-Western lens	8	Agarwal, Huang and Clark (2024) ; Daniel and Suryawanshi (2020) ; Chu, Li and Wang (2024) ; Idris, Fildzah and Basarah (2024) ; Smith (1999) ; Prabhakaran et al. (2022) ; Knox (2020) ; Rogoff (2003)

Lens/Topic	No of Papers (Approx.)	References
Equity/ digital divide	8	CBSE (2022) ; Central Board of Secondary Education (2022) ; Ministry of Education (2022) ; NCERT (2021); Schleicher (2018) ; UNESCO (2022, 2023) ; NITI Aayog (2018)
Teacher Collaboration and Design	5	DiSalvo (2009) ; Sanghvi (2024) ; Knox (2020) ; Sarpong and Gupta (2020); Selwyn and Facer (2013)
Affective Computing / Emotional Aspects	3	Kapoor, Burleson and Picard (2007) ; Holmes et al. (2019) ; Khuman (2024)
History and Philosophy of Education	4	Dewey (1933) ; Foucault (1977) ; Rogoff (2003) ; Schleicher (2018)
STEM and Language Learning	2	Daniel and Suryawanshi (2020) ; Chen et al. (2020)
Assessment and Academic Outcomes	3	CBSE (2022 reports) ; NCERT (2021); Facione (2015)
Policy and Strategy	5	Ministry of Education (2020, 2021, 2022) ; NITI Aayog (2018); UNESCO (2021, 2022)

9 USE OF POSITIONALITY

The authors are faculty members at Delhi University. This has helped them position themselves at a conceptual depth within varied relational contexts in Indian classroom realities. The insider's perspective has provided a

profound conceptual exploration with insight into how students in Indian classrooms respond to AI tools; how the pedagogical shifts are perceived within the Indian context, both by the teachers and the students. The paper follows a systematic representation of literature with a ‘reflexivity’ (Reflexivity’ is a systematic framework that is often used in a qualitative approach to research where the researcher is aware of their background and academic biases. He is under continuous self-examination throughout the process of research and can express her beliefs, backgrounds, and biases within the script of her research) framework in arguments. Contributions within the paper aim to build a new framework for discussing the potential impacts of AI-enabled education and suggest an outline to bridge the gaps in current discourse. The paper underlines the significance of the development of critical minds instead of Algorithmic minds, which further expresses itself in students’ agency and epistemic Delimitations

First, this paper presents a conceptual framework that can be further explored within empirical research. Hence, the present paper does not offer empirical generalisability. Second, the research positions itself in the Indian context with an in-depth understanding. It draws on global comparisons to illustrate best practices due to the paucity of research within the Indian context, with a critical framework.

10 DISCUSSION

‘AI algorithmic minds’ as a construct has been used to define the learner who may develop passive learning habits due to over-reliance on external feedback rather than their judgments. Such learners are not able to engage with critical and reflective questions that are significant to understand social, cultural, and linguistic biases. Such thinking can lead to diminished power that becomes the precursor for social change.

Algorithmic minds as an idea can be understood by drawing on the frameworks of Vygotsky's socio-cultural theory (1978), Bandura's concept of self-efficacy (1997), Foucault's theory of disciplinary power (1977), and Freire's critical pedagogy (1970). The socio-cultural theoretical perspective (Vygotsky) will be dealt with in the next section.

[Bandura \(1997\)](#) defined ‘self-efficacy’ as an individual's belief in their capabilities to be able to manage and organise the course of their lives. This belief system influences the way people approach their goals. Feedback given to the learners impacts their self-efficacy. Algorithmic feedback may lead towards an external locus of control in learners. This algorithm altogether neglects the personal turns that a student's journey can take due to extraordinary motivational factors such as good peer support, teacher's belief and encouragement, resilience built through failures, personal triumph in adversity, etc.

[Foucault \(1977\)](#) introduced the construct of ‘disciplinary power’ in his seminal work *Discipline and Punish*. According to him, there are pervasive and subtle forms of power that operate by constant observation, evaluation, and normalisation of the behaviour. In a system where there is constant surveillance, individuals gradually internalise such a control system and condition themselves to take it normally. They begin to regulate themselves accordingly. Foucault suggested that it works like a ‘Panopticon’ prison, where people are always visible to a watchtower, but cannot tell when they are being watched. This results in self-discipline according to the rules set by the system. Bringing this idea into today's educational landscape, one can witness that ‘AI-powered EdTech tools’ function in a similar way. These systems track students’ responses, failures, screen time, facial expression, attention span, and rank the students into categories based on their performance. Just like Foucault's Panopticon, when students have this idea back of their mind that every move of theirs will be evaluated, they might feel pressured. They might end up choosing options for their education journey where AI Algorithms support them instead of taking personal decisions based on real-life circumstances. This is kind of penalizing deviations, often ignoring individuality and the complexity of social-emotional factors. Thus, one can say that AI tools not just shape the minds based on algorithms but sort of control them under behavioural predictions and eliminate the possibility of creativity. [Freire \(1970\)](#) advocates reflective and liberating education. AI-powered algorithmic minds go exactly against the idea of freedom of mind and their individuality.

11 SOCIO-CULTURAL LEARNING THEORIES AND AI

The third pillar of this literature review comes from the socio-cultural theories of learning. [Vygotsky \(1978\)](#) sees cognitive development as a process mediated by social interaction and cultural tools. AI is the new tool that now mediates learning. The mediation is highly data-dependent and scores based on popular responses, which have been termed as AI algorithmic minds in the present paper. The question that needs deliberation is whether such a tool supports higher-order thinking.

Another concept that was proposed by Vygotsky is the Zone of Proximal Development (ZPD), which can be seen as relevant in the present paper. This concept explains that learners learn by interacting with a more knowledgeable other, be it the teacher, peers, mentor, parents, or any other person, depending on the context. In this case, learning is not passive absorption; rather, it is active involvement in social interaction, internalization of concepts through discussion, and developing understanding. There is a significant role of ‘scaffolding’ that helps learners grow through this process. Scaffolds are guided instructions that help learners to construct knowledge. AI, as it currently exists, struggles with these abilities of scaffolding. It provides convenient answers and delivers content. It does not tailor its responses to construct the understanding of the learner based on her socio-cultural context. It is significant that, in contrast, humans can evolve their understanding based on evolving contexts and tailor their responses according to the diverse groups of learners they are engaging with. This can become a significant shortcoming of AI in a multi-lingual and multi-cultural classroom. According to the UDISE+ (2021-22) report, vernacular language is a major medium of instruction in almost 60% of the schools ([Ministry of Education, 2022](#)). Artificial intelligence majorly uses English as the medium of engagement, with very little cultural embeddedness. Such systems can reinforce Western dominance over knowledge while marginalising indigenous, collective, and community-based knowing that is inherently diverse and contextual.

Artificial Intelligence places knowledge as a fixed output instead of co-creating it through constructivism, further risking the passive involvement of the learner. For example, Artificial intelligence serves the learner with

the correct answer, while taking away the opportunity to reach that correct answer by reflection. Moreover, the ‘why’ element is often completely neglected when learners are presented with ready-made answers. There are significant pedagogical and ethical questions that we need to ask as educators before being blinded by technological enthusiasm.

12 RETHINKING THE ALGORITHMIC TURN IN INDIAN EDUCATION

According to the UDISE+ (2022) report, India is witnessing a surge of 70% in the adoption of digital tools by secondary schools. There is a discourse around celebrating efficiency, speed, and data-driven thinking. However, the critical question is what AI is doing to the developing minds in our classrooms. One has to rethink the difference between machines and humans. Successful learning signifies developing autonomous minds with curiosity and a reflective attitude. The shift towards AI-driven integration is creating ‘followers of data’ who base their decisions not on reflection but on scores. There is a rapid search for correct answers, then engaging with the process to be able to think about the correct answer. Performance matters, and participation is neglected. This creates a ‘shallow mind’ that is never challenged and is conveniently out of the fear of being seen or questioned.

The construct of ‘AI algorithmic minds’ poses not just a ‘technological critique’; rather, a serious ‘pedagogical critique’. Learners need freedom of mind and contextual thinking. Scores and data can help them reach majoritarian perspectives; however, developing the critical mind that can think beyond numbers is the real purpose of education. Learners are not robots; they need to shape their reality. They need to be the thoughtful minds without losing the critical touch. Moreover, seeking quick answers can lead to the development fear of ambiguity.

Paulo [Freire's \(1970\)](#) criticism of the ‘banking model of education’ becomes an important theoretical perspective to understand this situation. The phenomenon elucidated an educational context where knowledge is deposited and treated like a bank commodity, rather than participating,

processing and constructing. AI looks at the knowledge in a similar manner. Knowledge is something that can be gathered or collected like money instead of being constructed. Learners in such a scenario become collectors of knowledge (earlier from textbooks, now from AI-driven tools) instead of gaining skills to construct knowledge for a lifetime. Data conformity over reflection, insight, and lived experiences may create a set of learners who undervalue social, cultural, language, and regional disparities.

13 THE INVISIBILITY OF CONTEXT AND THE DISAPPEARANCE OF ‘WHY’

AI-driven tools are not equipped to deal with context-based reality. They cannot deal with regional contexts, community diversity, or multilingual realities. They speak in neutral tones. Neutral and standardised solutions are not effective problem solvers. Rather, this shows a kind of cultural bias that further results in the marginalisation of viewpoints that are not mainstream. There are local and indigenous worldviews that need integration while teaching- learning transactions. The historical struggle of several eminent educationists who have put their lives to create a reflective and critical pedagogy to create minds that can ask ‘why’ at the right place is deliberately neutralised. [UNESCO \(2023\)](#) report depicts that AI-driven platforms seldom go beyond translation; even fewer platforms engage with regional/local viewpoints. Whatever is showcased as the local viewpoints is not meaningful. For example, a student might be seeking knowledge on some physics phenomenon that occurs locally. AI might miss stories around how such occurrence of the physics phenomena was conceptualised by earlier generations. Such local ‘oral memory’ of engaging with the physics phenomena may be processed while collecting real-life data through inter-generational stories. Such stories may help the learner to locate themselves in local sources of knowledge and may give insights into how to draw logical, scientific, and systematic insights from such stories. The entire process is significant for the learner to develop their sense of physics.

14 AI-DRIVEN COMMUNICATION

A growing body of research in the Indian education system reveals that students are increasingly feeling disengaged without having real mentorship to motivate them. AI ‘virtual mentors’ are gradually becoming popular in use in Indian schools following CBSE and NEP guidelines. These systems offer personalised feedback, recommendations, a support system, doubt assistance, and revision requirements ([Bardia and Agrawal, 2025](#); SSRVM blog, 2025). This seems particularly beneficial as it reduces the requirement for human attention ([Sanghvi, 2024](#)), especially in larger under-resourced classrooms. Large-scale mentor programs benefit from the use of AI as they provide personalised tracking of mentees’ needs without real-time tracking, such as AICTE’s ‘One Student One Mentor’ program (Best Colleges, 2025). What is blatantly ignored is the human touch in every problem that ensures personal understanding and a feeling of supportive community, by which students get motivated and feel safe.

AI tools like Quizizz, Labster, and other educational games are also becoming popular amongst parents who want their children to be busy during their working hours. According to [Sanghvi \(2024\)](#), AI tools can be used to further boost participation across diverse settings like rural India. Importantly, all these reports and articles don’t fail to mention that AI cannot replace teachers; it is just to enhance their capacity. However, the picture is not as black and white as it is being portrayed.

This is to understand that when an algorithm takes over the role of a guide, nuanced emotions and incidental distinctions are missed, such as body language, cultural references, or even silence. These are an essential part of Pedagogical traditions. The non-verbal communication, including facial expressions, body language, gestures, and deliberate/non-deliberate silence, carries considerable meaning while teaching/mentorship. In Indian culture, especially communicative practices draw heavily from subtle culturally coded cues of communication, which cannot be understood by AI ([Prabhakaran et al., 2022](#); Idris, N. S., Fildzah, H. N., and Basarah, M. F., 2024; [Agarwal et al., 2024](#)). Cultural forms of communication are fundamental to Indian diverse traditions ([IJARSCT, 2024](#)). Ignoring such forms can result in alienisation of a large scale of population that uses these

cues while communicating. Moreover, there is a huge risk of losing the rich Indian oral traditional pedagogical methods such as storytelling and dialectical learning.

Humans, while interpreting silence and non-verbal cues, can adapt to the communication, specifically connecting to the socio-cultural background of the learner. They can show emotional intelligence intertwined with cultural meanings that are critical aspects for sustained and meaningful talk. These roles cannot be played by algorithmic guides.

15 SURVEILLANCE AS A GROWING CULTURE

AI tools now have built-in Eye-tracking systems and emotional analytics. They analyse every keystroke of the learner with a camera constantly tracking them. This is similar to Foucault's 'Panopticon metaphor' (1977) that ensures constant surveillance over the learner. Learners are now performing for the algorithm. This is not learning that is constructed out of curiosity. Future learning is then compromised as the real curiosity, which would have set the path for it has not been generated. Already education system grapples with exam-based performance anxiety, and worldwide, efforts have been made to replace it with a genuine discovery and learning cycle, which gets defeated by AI-driven surveillance. Second, one has to understand that learning is a complex process with a mix of varied internal thoughts that do not necessarily get transferred into behavioural outputs.

16 RETHINKING 'EFFICIENCY' AS A PEDAGOGICAL GOAL

As an educational thinker, this is a question to be asked: Do 'efficiency' and 'speed' matter? Machines have already been doing their work efficiently for ages. Then, where do we need the humans? We need humans to redirect the machine thinking towards contextual conditions and to redevelop technology for the newer, upcoming needs. Educational goals for learners are to gain perspectives and develop critical thinking skills. Education is not

quick; it is relational, messy, and often contradictory. A teacher's narratives that are not fixed and keep developing over a period of time, and according to the context, are an essential part of this process. Generating answers efficiently and quickly is the product, but this process is missing out on this journey. Hence, replacing teachers is not an alternative ([Schleicher, 2018](#)); rather, it is significant that we understand digitalization and technology-driven tools are to be integrated with much thought. We cannot get stormed with their blind applications.

Researchers have found a significant correlation between digital dependence and the decline of metacognition. Risko and Gilbert (2016) describe that when thinking is continuously delegated, it manifests into a habit generation within students where they habitually do not extend their thinking to get deeper into the concepts. They describe it as ‘cognitive offloading’. Indian government reports are also taking cognizance of this situation. CBSE Annual Report (2022) mentions that with AI, students now have reduced depth of engagement. This results in difficulty in the application of concepts in a variety of fields. Over-reliance on machines to think, assess, and suggest has created a sort of mundane thinking pattern. It has the potential to impact their self-regulated learning, which, according to [Flavell \(1979\)](#), is one important skill for metacognition (Metacognition is a psychological term that is widely used in the context of education. It means ‘thinking about thinking’. A skill that is required to maintain self-regulation and awareness over one's thinking skills to manage their journey of learning) in students. Students are then less likely to engage with self-planning and monitoring.

17 STUDENTS' EPISTEMIC SELVES AND TECHNOLOGICAL CONTEXTS

Learners, during their educational experience, shape their selves that helps them take a position in the world outside. Researchers have pointed out that identity formation in educational settings is formed by various factors such as feedback, peer interactions, and a narrative that students construct about their abilities ([Yeager and Dweck, 2012](#)). Till now, human intervention on

all the mentioned stages has added a reflective and critical touch, which is required to embed the ‘context-sensitivity’ in the learning environments of the classrooms. However, we need to ask ourselves- how could an algorithmic feedback and interactive system create thoughtful minds?

When an AI system analyses data, it scores and ranks on the basis of students’ potential and capabilities. Over time, these messages are internalised by the students and impact their academic/epistemic self-concept. They further impact the motivation and aspirations of the students, either encouraging or limiting them on the basis of their earlier performance. This is where the idea of “algorithmic minds” is particularly significant.

18 GLOBAL SOUTH COMPARISONS

Similar broader patterns can be observed in countries like South Africa, Nigeria, and Brazil in their integration of AI-driven technologies in education. In Nigeria, AI-assisted programs have shown promising results despite infrastructural limitations and restricted resources, highlighting their potential as a tool for education. Brazil is using AI-driven technologies to extend education into remote areas ([Penprase, 2024](#)). However, with them also, the restriction of resources creates a digital divide. Integration of local culture remains a pertinent question, and language barriers further complicate the situation. This is because most AI tools initially adopt the most dominant language, and a gradual and sporadic movement is made in other languages, which also do not justify the linguistic and dialectical diversities. This further results in sidelining indigenous perspectives. Remote regions in South Africa and Latin America underscore the need for culturally relevant content ([Penprase, 2024](#)).

19 WAY FORWARD: INTRODUCING THE PACE+ MODEL

The question that is pertinent to ask is whether we can integrate AI in education or not. The answer is complex as opposed to the question. First,

the goal is not to integrate, but to humanise the process. A culturally sensitive and emotionally intelligent outlook is needed in this process. This integration needs to make the learner's experience the centre. Automation has to take the back seat in some places. This paper invokes a model to support the thinking process of the people who will be doing this integration. The model has 6 pillars- participation, awareness, contextualisation, engagement, empathy, and slowness (PACE+ model). This is a shift that needs to be integrated for AI-bots that are used for educational purposes. It is like reimaging AI use in learning, and not just producing speed and efficiency as the only skill to offer to our learners.

‘Participation’ means, co-creation of AI tools with educators. People who are engaged in direct interaction with the students on a grassroot level. This will help in involving local viewpoints, indigenous meanings, and interpretations that are unheard, in turn supporting the cultural alignment ([Selwyn and Facer, 2013](#); [DiSalvo, 2009](#)). This is all the more needed in countries like India, where there is an abundance of local, regional, indigenous, and disempowered voices. ‘Awareness’ in the context of this paper means a deeper and more critical understanding of how AI works for both teachers and students. The people in the education system need to engage with how AI shapes learning. This will create awareness around when to question AI data and suggestions. This has been described as ‘algorithmic awareness’ by [Williamson and Eynon \(2020\)](#). The researchers describe it as an essential skill that can help learners and educators avoid trusting AI blindly. In India, the government has recognised the importance of this dimension. The National Digital Education Architecture (NDEAR) talks about digital literacy as the need of the hour. Here, ‘digital literacy’ means critically engaging with how AI interferes with and supports our thinking process. ‘Digital literacy’ as a terminology in the context of education has been coined by [Belshaw \(2011\)](#) in the overall context of the digitalisation of educational tools. According to Belshaw, digital literacy includes cultural and critical knowledge embeddedness. The idea of ‘critical digital literacy’ has also been supported by Kelner and Share (2007), who argue that learners are to engage with media and technology thoughtfully. Practically, this means that teachers need to go beyond- how to give a prompt, how to click and code. It means helping students to ask: what were the recommendations made, who benefited from these recommendations,

whose standpoint does it support, and who are getting unheard by these suggestions? Learners cannot be passive in any of the educational spaces.

‘Contextualisation’ means being adaptive to the rich multicultural, and multilingual settings of any society. Standardised systems do not engage with local realities, folk knowledge, and indigenous pedagogies effectively. It is a sort of ‘epistemic injustice’ with the students when folk traditions and knowledge systems are pushed to the periphery ([Smith, 1999](#)). [UNESCO 2023](#) report, while recognizing this concern, stated that all the Edtech tools must take cognizance of this gap. [Daniel and Suryawanshi \(2020\)](#), while researching multilingual AI tools, witnessed the transformative impact of these on Indian students.

‘Engagement’ of students is another chord that has to be retained while working with the tools of AI. We are talking of developing a tool for the students, and hence, their engagement on every level in the form of suggestions must be incorporated. Anyways, education is a relational context where there is a dialectical relationship between the teacher and the student. Knowledge is co-created, and AI should enhance, not replace, such maxims.

A separate AI-driven technology can be introduced for educational purposes that does not typically produce direct answers, rather, it helps in reaching probable answers by providing relevant scaffolds to the learner. The process by which one can reach towards answer should be made a part of this conversation. It is to state that even if AI is developed to hold such conversations with the learner that scaffold appropriately, it cannot replace the pedagogical discussion a teacher can hold within a classroom. This is because the range with which a teacher can engage in a meaningful discussion with the learners, keeping in mind their diverse socio-cultural backgrounds, emotional journeys, and motivational aspects, is far broader and more nuanced than what even most of the AI can achieve.

20 THE ‘+’ SIGN IN PACE+ INTRODUCES TWO MORE SIGNIFICANT CONCEPTS: EMPATHY AND SLOWNESS.

‘Empathy’ here practically means being able to deal with the emotional aspects that one can judge with the conversational aspects of the learners. AI needs built-in technology that can judge confusion, frustration, and anxiety in learners. Emotionally aware tutoring systems improve the motivation for learning ([Graesser *et al.*, 2014](#); [Kapoor *et al.*, 2007](#)). For developing learners, when their traits are still in the making, such dispositions in AI-driven interfaces are essential.

‘Slowness’ signifies a pedagogical resistance to haste for quick output and efficiency metrics. A little slowness and wait time allow learners to grapple with complexity and come up with skills and solutions in the form of learning ([Berg and Seeber, 2016](#)). Pioneers in educational theory, such as [Vygotsky \(1978\)](#) and [Dewey \(1933\)](#), have argued for the construction of knowledge for deep engagement. There is a requirement to design AI tools that support iteration and process-oriented learning.

For example, in a practical setting of a rural school, local teachers and community workers collaboratively work to create storytelling modules that are based on local knowledge and the history of water conservation in the area. Students first connect with the concepts and examples in the classroom, followed by real-life data collection and using AI to generate multimedia presentations. Through this journey, they also come up with the knowledge of how similar problems have been tackled in other parts of the world and critically engage to conduct a synthesis of their understanding, where they come up with creative ideas to handle the problems. The solutions are contextual and adaptive to the multi-cultural environment, incorporating local needs. Further group discussions are initiated that help in the development of empathy on part of the students as they are engaged with lived experiences. Slowness is practiced by the teachers to let the process of learning unfold without hastening it with the help of AI. Instead, they plan and use AI in a way that augments the learning process rather than hindering its usual course.

REFERENCES

- Agarwal, V., Huang, J., and Clark, J. (2024). AI writing assistance and the homogenization of cultural voice in global south scholarship. *arXiv preprint arXiv:2409.11360*. <https://arxiv.org/abs/2409.11360>
- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman.
- Bardia, A., and Agrawal, A. (2025). MindCraft: Revolutionizing Education through AI-Powered Personalized Learning and Mentorship for Rural India.
- Banerjee, A., and Biswas, R. (2025). Empowering educators, bridging divides: Evaluating Diksha and Swayam for digital teacher training and economic inclusion in India. *International Journal of Creative Research Thoughts (IJCRT)*, 13(7), e710–e726
- Berg, M., and Seeber, B.K. (2016). *The Slow Professor: Challenging the Culture Of Speed in The Academy*. University of Toronto Press.
- BestColleges. (2025). Rebooting the gurukul system: Can AI mentor India's next million engineers? *BestColleges India Reports*.
- Belshaw, D. (2011). *The Essential Elements of Digital Literacies* (Doctoral dissertation, Durham University). Durham E-Theses. <http://etheses.dur.ac.uk/3446/>
- CBSE. (2022). *CBSE annual report 2021–22*. Central Board of Secondary Education.
- Central Board of Secondary Education. (2022). CBSE introduces Artificial Intelligence as a subject in schools. *Ministry of Education, Government of India*. <https://www.cbse.gov.in/>
- Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chu, C.H., Li, R., and Wang, Y. (2024). Cross-Cultural Differences in Using Nonverbal Behaviors to Identify Indirect Replies. *International Journal of Intercultural Relations*, 95, 65–78. <https://doi.org/10.1016/j.ijintrel.2024.01.002>
- Daniel, J., and Suryawanshi, S. (2020). AI and multilingualism: Leveraging language technologies for inclusive education in India. *Journal of Educational Technology*, 17(3), 35–47.
- Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. D.C. Heath and

Company.📄

DiSalvo, C. (2009). Design and the construction of publics. *Design Issues*, 25(1), 48–63.📄

ET Edge Insights. (2024). AI in School Education Transforming Indian Classrooms.📄

Facione, P.A. (2015). Critical thinking: What it is and why it counts. *Insight Assessment*.📄

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906–911.📄

Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison* (A. Sheridan, Trans.). Pantheon Books.📄

Freire, P. (1970). *Pedagogy of the Oppressed*. Herder and Herder.📄

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.📄

Graesser, A.C., Chipman, P., Haynes, B.C., and Olney, A. (2014). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, 48(4), 612–618.📄

Hatkar, N. (2025, May 19). AI in education: How technology is shaping the future. 21K School. <https://www.21kschool.com/in/blog/ai-in-education/>.📄

Holmes, W., Bialik, M., and Fadel, C. (2019). *Artificial intelligence in education: promises and implications for teaching and learning*. Center for Curriculum Redesign. <http://curriculumredesign.org/wp-content/uploads/AI-in-Education-Promises-and-Implications.pdf>📄

Idris, N.S., Fildzah, H.N., and Basarah, M.F. (2024). Patterns of verbal and nonverbal communication of Indian learners in BIPA learning. *Atlantis Press*.📄

IJARST. (2024). Impact of Artificial Intelligence (AI) in Education on Student Learning and Teacher Functionality.📄

IndiaAI. (2024). How to harness AI in schools: Opportunities and challenges. *IndiaAI Research Review*.📄

India Today. (2025). How artificial intelligence is reshaping personalised learning in India's schools. *India Today Education Special Report*.📄

Jain, B. (2025, May 30). PM e-Vidya: Digital education for inclusive learning 2020. *IMPRI Impact and Policy Research Institute*.

<https://www.impriindia.com/insights/pm-e-vidya-digital-education-for-2020/>

- Kapoor, A., Burleson, W., and Picard, R.W. (2007). Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65(8), 724–736.
- Kellner, D., and Share, J. (2007). Critical media literacy, democracy, and the reconstruction of education. In D. Macedo and S.R. Steinberg (Eds.), *Media Literacy: A Reader* (pp. 3–23).
- Khan, R., Qamar, M.T., Ansari, M.S., and Yasmeen, J. (2025). Enhancing or impairing? Exploring Indian EFL learners' academic writing narratives with ChatGPT. *Cogent Education*, 12(1), 2514329. <https://doi.org/10.1080/2331186X.2025.2514329>
- Kincheloe, J.L. (2005). *Critical constructivism*. Peter Lang.
- Knox, J. (2020). Artificial Intelligence and Education in China. *Learning, Media and Technology*, 45(3), 298–311. <https://doi.org/10.1080/17439884.2020.1745403>
- Luckin, R., Holmes, W., Griffiths, M., and Forcier, L.B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education. [https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/Intelligence Unleashed Pub 2016.pdf](https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/Intelligence%20Unleashed%20Pub%202016.pdf)
- Mehta, A.C. (2024, July 12). AI in Indian school education: Challenges, curriculum, & teacher training. Education for All in India.
- Ministry of Education. (2020). *National Education Policy 2020*. Government of India.
- Ministry of Education. (2021). National Digital Education Architecture (NDEAR) Blueprint. Government of India.
- Ministry of Education. (2022). *PM eVidya Programme*. Government of India. <https://www.education.gov.in/>
- Ministry of Education. (2022). Unified District Information System for Education Plus (UDISE+) 2021–22 Report. Government of India. <https://udiseplus.gov.in/>
- Ministry of Education, Government of India. (2022). UDISE+ Report 2021–22. <https://udiseplus.gov.in/#/home>
- National Digital Education Architecture (NDEAR). (2022, November 11). Open standards & specifications for NDEAR. Ministry of Education,

- Government of India. https://www.ndear.gov.in/images/pdf/NDEAR-Open%20Standards_Version.pdf
- NCERT. (2021). *Learning Outcomes at the Elementary Stage*. National Council of Educational Research and Training, Government of India.
- NITI Aayog. (2018). *National Strategy for Artificial Intelligence*. Government of India. [↵](#)
- Penprase, B. (2024, September 24). *The AI Education Leapfrog in the GLOBAL SOUTH*. *Forbes*. <https://www.forbes.com/sites/bryanpenprase/2024/09/24/the-ai-education-leapfrog-in-the-global-south/> [↵](#)
- Prabhakaran, V., Kaur, S., Jernite, Y., and Dodge, J. (2022). *Survey of cultural awareness in language models: Text and beyond*. arXiv preprint arXiv:2211.13069. <https://arxiv.org/abs/2211.13069> [↵](#)
- Risco, E.F., and Gilbert, S.J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Rocco, T.S., and Plakhotnik, M.S. (2009). Literature reviews, conceptual frameworks, and theoretical frameworks: Terms, functions, and distinctions. *Human Resource Development Review*, 8(1), 120–130. <https://doi.org/10.1177/1534484309332617>
- Rogoff, B. (2003). *The Cultural Nature of Human Development*. Oxford University Press. [↵](#)
- Sanghvi, K. (2024). Impact of artificial intelligence in secondary classrooms: Enhancing student engagement and mentorship. [↵](#)
- Schleicher, A. (2018). *World class: How to build a 21st-century school system*. OECD Publishing. [↵](#)
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.
- Selwyn, N., and Facer, K. (2013). *The politics of education and technology*. Palgrave Macmillan. [↵](#)
- Siddiq, I., Selvin, S., Shobana, D., and Suresh Kumar, J. (2024). Challenges and opportunities of AI implementation in the education systems of rural India. Council for the Advancement of Peer-Reviewed Journals. <http://capdr.org/wp-content/uploads/2024/12/47.-Dr-I.-Siddiq-et-al.pdf> [↵](#)

- Smith, L.T. (1999). *Decolonizing methodologies: Research and Indigenous peoples*. Zed Books. [↗](#)
- SSRVM. (2025). AI tutors and virtual mentors: The future of personalized learning in Indian schools. *SSRVM Blog*. Retrieved from SSRVM website. [↗](#)
- Tarde, Y., and Joshi, P. (2023). A study of effect of cognitive offloading on instant performance and metamemory in short-term task. *Indian Journal of Behavioural Sciences*, 26(2), 85–94. <https://doi.org/10.55229/ijbs.v26i2.03>
- UNESCO. (2021). Artificial intelligence and education: Guidance for policy-makers. *United Nations Educational, Scientific and Cultural Organization*. <https://unesdoc.unesco.org/ark:/48223/pf0000367823> [↗](#)
- UNESCO. (2022). 2021/2 *Global Education Monitoring Report: Non-state Actors in Education: Who chooses? Who loses?* United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000379875> [↗](#)
- UNESCO. (2023). *Global Education Monitoring Report: Technology in Education*. Paris: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000384551> [↗](#)
- Vygotsky, L.S. (1978). *Mind in society: The Development of Higher Psychological Processes*. Harvard University Press. [↗](#)
- Williamson, B., and Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235. <https://doi.org/10.1080/17439884.2020.1798995> [↗](#)
- Xia, Y., Huang, C., Duan, Y., and Wang, T. (2022). A comprehensive review on artificial intelligence in education. *Education and Information Technologies*, 27(1), 829–857. <https://doi.org/10.1007/s10639-021-10717-9> [↗](#)
- Yeager, D.S., and Dweck, C.S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47(4), 302–314. <https://doi.org/10.1080/00461520.2012.722805> [↗](#)
- Zawacki-Richter, O., Marín, V.I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of*

Educational Technology in Higher Education, 16(1), 1–27.
<https://doi.org/10.1186/s41239-019-0171-0>

Artificial Intelligence and Crime: The Potential Challenges of the 21st Century Society

Ayush Raj and Ananya

University of Delhi

DOI: [10.1201/9781042018116-19](https://doi.org/10.1201/9781042018116-19)

ABSTRACT: Marching forward from the Golden Age of the British Empire to the 21st century multipolar world, mankind has travelled way ahead in terms of scientific advancements, defence technologies, space exploration and the most recent development i.e., Artificial Intelligence. Since the last hundred years, if anything that is so discussed, debated after the invention of nuclear technology is the advancement of artificial intelligence. AI has a potential to transform the society entirely but it comes with some of the most dangerous challenges that could arise if it is not strictly regulated. Advancement in AI and vast amount of online data, available just over the click, amplifying criminal abilities which requires modern solution over traditional human intelligence. This paper will trace the potential challenges of AI in our society. The specific focus will be on AI-driven crimes in our society such as Deepfakes, AI surveillance, and Phishing. The paper will search for probable reasons how AI can be a potential threat in this post-modern world! The objective of this paper will be to make people aware about the significant problems that AI can cause in our everyday life. Case studies are also dealt in details about the discussed crimes, along with measures that can reduce the rate of perpetrated crime.

Keywords: Artificial Intelligence, crime, surveillance, deepfakes, and cybercrimes

1 METHODOLOGY

This paper adopts a qualitative, descriptive–analytical methodology combining doctrinal analysis with illustrative case studies. Secondary sources—including legal frameworks, policy reports, crime statistics, and documented cases of AI-enabled crimes such as deepfakes, surveillance misuse, and phishing—are systematically reviewed. A problem-oriented analytical framework is applied to identify emerging risks, assess societal implications, and evaluate regulatory and preventive measures for mitigating AI-driven criminal threats.

2 INTRODUCTION

About more than a century ago when Britain was involved in the First World War (1914-1918), it was affected the most. On the economic front it was facing stagnation due to fighting on multiple fronts, such as defending its borders with Irish hostility. After the partition of Ireland from Great Britain in 1921, there was relative peace and stability. The British society was also churning itself. The class hierarchy was diminishing slowly as Gentry class was gradually losing its foothold, as source of income for them decreased. They were either selling or renting their estates and land in order to generate income out of it. Historian Arthur Marwick in *The Deluge* talks about the transformation that was happening after the Great War. “Marwick says that class distinctions softened, national cohesion increased, and British society became more equal during the period” (Marwick, 1965, as cited in Wikipedia, 2025). There was a sharp rise in investment for new technologies such as automobile making for better transportation, military equipment, creation of new power looms. Printing press were increasing that led to diversification of rational voices against the war jingoism. Siegfried Sassoon, one of the most spoken anti-war poets, also served in the

war, wrote an open letter of protest to the war department. He termed it as a deliberate attempt to prolong the war. In *Trench Duty*, Sassoon laments the ferocity of war,

Shaken from sleep, and numbered and scarce awake,
Out in the trench with three hours' watch to take,
Hark! There's the big bombardment on our right
Rumbling and bumping; and dark's a glare
Of flickering horror in the sectors where
We raid the Boche; men waiting, stiff and chilled,
Blank stars. I'm wide-awake; and some chap's dead. ([Sassoon, 1917](#), p. 123)

3 GOLDEN AGE OF DETECTIVE FICTION

In England generally the period between the 1920s and 1930s is considered as Golden Age of Detective Fiction. Between the two world wars, this period saw numerous transformations as in economy, scientific temperament, perspective towards wars were also changing. Common masses were now fed-up of violence after the First World War. Instead of glorifying wars they were now seeing it as burden, that exhausts resources and human life. People of that time were tired of reading emotional works and now new lease of fresh air was required to balance the emotional sabotages that the War gave. People were in want of storyline that hit the neurals. Through the plots of crime fiction, writers were able to gather public attention and sentiments for victims, and respect for the detective, police. The era saw the rise of classic detective stories, cosy mysteries, often characterised by tight and complex plots, intricate puzzle that runs from start to end, clever and stud detective, whose focus was on solving the case and finding the criminal. It shows the fair play where the readers should be able to solve the cases using the information presented in the novel, without relying on the information presented. It contributes to the entertainment and maintain credibility among readers. Central element of Golden Age is “Whodunnit” where the readers are allowed to think critically and identify the perpetrators of the crime based on the collected

evidences through logical deductions. The ‘novels of sensation’, ‘vendors of excitement’, ‘preachers to nerves’ as coined by Colin Watson for the phenomenal intellect who lodged in an unfashionable London street, searching for the criminal (Ch. 1). The most prominent figure who rose to stardom was Sherlock Holmes created by Sir Arthur Conan Doyle. Resident of 221B Baker Street, London, Holmes was a private detective of great intellect along with his friend Dr. Watson. His extraordinary zeal and proactive mind were fuselage for the investigations he performed. His deductions out of any conversation or crime scenes are like a telepathy, where all truth are converged one after the other. The London that Doyle creates is like a myth filled with imagination. Watson further comments, “it is quaint enough to make nostalgic appeal. It is exceedingly well-ordered: those telegraphic offices are never closed, no cabs is ever otherwise than within instant hail” ([Watson, 1971](#)). Doyle managed to manipulate the minds of readers that Holmes is one-stop solution for every crime that is committed in and around London. “It is a city whose every crime is soluble and whose vices are sealed within narrow and defined areas. It is a cosy place. It is, for as long as a hawk-eyed man broods in Baker Street, a safe place” ([Watson, 1971](#)).

Agatha Christie, often known as Queen of Crime, known for her extensive writing on detective fiction. She created two most popular detectives, one male Hercule Poirot and another female amateur detective Miss Jane Marple. Setting of the stories were country mansions, exotic beaches, quaint villages and brimming room with secrets. Miss Marple was Christie's mature brat, up against the evils of society. She is a detective that relies on her intuition and her deep understanding of human psychology puts her ahead than federal police. She is not armed with gadgets, advanced technologies, instead decodes psychology of people while solving crime, when technology played no significant role. She observes minute details and uses her knowledge of human nature, gained from years of experience living in close quarters of society. AI also functions in that way where it analyses the vast amount of data and algorithms thereby identifying patterns and motives that humans might miss and can predict crime that can happen in future. *BBC* in collaboration with Agatha Christie estate has instituted a revolutionary AI-generated writing course featuring a digital “resurrection” of the Miss Marple. Using AI enhancement and using archival digital

materials, the curriculum presents Christie as a virtual instructor teaching crime writing. There is a '6 Minute English' conversational series published in *BBC* titled "Can AI Solve Crime?" Several questions were asked regarding the police procedural on account of different techniques involved in investigation. "Solving crimes involves collecting and reviewing lots of evidence, like DNA, fingerprints, or CCTV" (2025). Regarding AI help in police procedural that could speed up the work and could even solve crime as well. Prof. Ruth Morgan, an expert in forensic science and crime "thinks AI could help speed up police work, as she told *BBC Radio 4's* The Artificial Human" (2025). AI can help us marking and then matching patterns with the available database and have real-time results. In the end the host asked can AI-powered robots would able to ever replace a detective anytime. The answer is no; it could help humans in the detection process but cannot directly intervene in the process.

4 DEEPPFAKES

In a most unexpected, although deeply disturbing visuals came from the New Zealand parliament where a member showed her nude images made by deepfake technology, broadly AI. MP Laura McClure bravely displayed her naked image in front of the public. It took her just five minutes to create it using AI and looks real. This was an eye-opener episode in the ruthless world of AI. Technology that can foster human intelligence, is now creating trouble for the people. It was a powerful message to the world illustrating the danger Deepfake possess. As the name suggests, Deepfakes are technology that is involved in creating a fake image, manipulative content that involves use of deep learning techniques to create real, believable content that has never happened in real life. Homeland Security termed it as "synthetic media" or "synthetic content". It defines it as "synthetic media as any media which has been created or modified through the use of artificial intelligence/machine learning (AI/ML), especially if done in an automated fashion" ([Increasing Threats of Deepfake Identities, n.d.](#)). For McClure, it is not the technology that is problematic but rather its misuse is a problem. Since its inception in 2017, there has been many developments in

Deepfakes and synthetic media contents. Some of the most important deepfake cases are, Jordon Peele as like Barack Obama, creating a convincing audio and video fakes. The ObamaPeele video that has surfaced over the internet has been created using FakeApp. It's the same free AI tool that is used in creating fake porn images and videos. It is a crime against humanity which disturbs the mental state of the traumatised victim. Many of such cases go unreported due to lack of complaints due to social stigma. Another such edited video of then American House Speaker Nansi Palosi was viral here she was seen intoxicated where her voice was slurred. In the original video, she was sipping water and here the speed of the video was slowed down so that she looks like drunk. Swapping or Face Swapping is a precursor to deepfake technology. It was done using photo editor software such as Adobe Photoshop. Today AI is acting more convincingly to create morphed images. The face of one person is overlapped with the body of another person, which looks perfectly real. It could be done using Encoder or Deep Neural Network (DNN). Some of the applications that allows face swap are FaceSwap, Deepface Lab, Reface, Faceshifter, TikTok, etc.

Deepfakes, videos intended to spread malice against the sensitive information pertaining to Indian elections were rampant during the 2024 general elections. It is just synonymous to fake news during the 2014 and 2019 elections. India has fully entered into the digital campaign era, where AI is the new toolkit which is used to defame, make false narrative against the rivals. Though it is entirely different from the morphed videos which uses digitally blending of images and voices in a seamless manner, where gradual transformation of one image or shape into the another. According to a report published in *Le Monde* “Dead politicians resurrected to support a candidate in Tamil Nadu; a Muslim party leader intoning Hindu devotional song; usually discreet Bollywood stars Ranveer Singh and Aamir Khan openly criticizing the Indian prime minister and lending their support to the Congress, the main opposition party” (Sophie Landrin, 2024). In a similar incident, Mumbai police has registered a case against the Maharashtra Youth Congress Social media handle for creating and circulating deepfake video of Union Home Minister Amit Shah, where he is seen announcing the curtailment of reservation rights of Schedule Castes, Schedule Tribes and the OBCs. While the original video is to uphold the reservation to all concerned stakeholders by quashing the Muslim reservation in Telangana.

This defamatory intention could possibly be one of the reasons for the BJP to not cross the 400 plus seats.

Steps are being taken to stop the spread of such malicious contents being shared online. Press release by the PIB Delhi on 4th April, 2025 shares the measures taken by the Ministry of Electronics & IT, Government of India, where under The Information Technology Act, 2000 (IT ACT) have designed legal framework to protect the internet from unlawful activities that will ensure trust and safety of citizens. Various punishments are listed under the law for identity theft, cheating by personation, depicting children and under-aged women in sexually explicated content, browsing child abuse contents, publishing/transmitting contents that are obscene, violation of privacy, etc. The Indian Computer Emergency Response Team (CERT-In) issues warnings, alerts, advisories related to cyber threats, malicious attacks using Artificial Intelligence. CERT-In operates the Cyber Swachhta Kendra (Botnet Cleaning and Malware Analysis Centre) which aims to detect malicious programs and provide cyber security tips for citizens. Ministry of Home Affairs (MHA) has established the Indian Cyber Crime Coordination Centre (I4C) to expedite matters of cybercrime by launching a dedicated website to report all types of cyber-crimes including financial fraud on National Cyber Crime Reporting Portal (<https://cybercrime.gov.in>). Dharmendra Tewari and Navin Sreejith in PIB article *Government of India Taking Measures to Tackle Deepfakes*, promulgates the views and response of Indian government towards the evolving crisis due to misuse of AI.

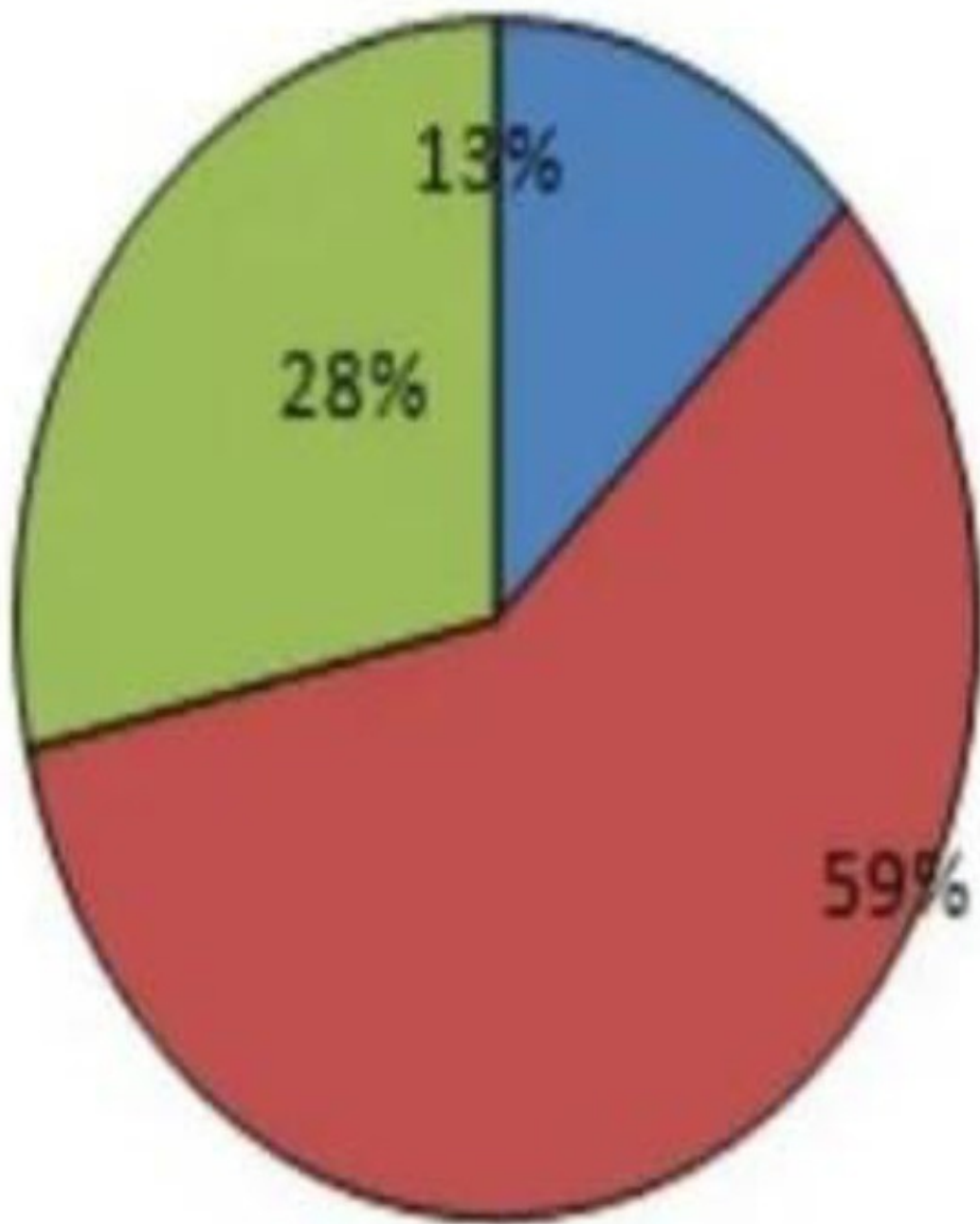
5 SURVEILLANCE AND PRIVACY VIOLATION

University College London's (UCL) ran a PanoptiCam project which tests, among other things, surveillance algorithms ([McMullan,2015](#)), to discuss contemporary ways of supervision, surveillance as given by Michel Foucault in *Power: Essential Works 1954-84*. Jeremy Bentham's idea of Panopticon, where he talks about the centrality of prison, where all prisoners are visible from the watch tower by the guard, but having limited visibility. This leads to the active argument of Michel Foucault through the developed concept of Panopticism, where it is an internalised surveillance.

An individual regulates its behaviour and itself, even though it may not be actively observed, assuming that they are constantly under surveillance. Foucault further elaborates that “one would allow a single gaze to scan the greatest number of faces, bodies, attitudes, in the greatest possible number of cells. The state presents itself as a certain spatial arrangement of individuals, in which all are subjected to a single surveillance” (1994, 72). Foucault cites Jean-Baptiste Treilhard, Councillor of State, Jurist of the Empire, in his *Code of Criminal Procedure of 1808* where he talks about the new role assigned to the prosecutor. His role is not only to prosecute the offender but also has a proactive role of supervising individual, so that he does not commit the crime. “The prosecutor must not have as his only function that of prosecuting individuals who have committed offenses; his main, primary function must be that of supervising individuals even before the infraction has been committed. A prosecutor is not just the agent of law who acts when law is violated; the prosecutor is, above all, a gaze, an eye constantly trained on the population” ([Foucault, 1994](#), 73). His primary argument of surveillance is that, at every level the information must be transmitted to the higher level which would ultimately result into sharing the information to the highest authority, here the emperor, watching the entire society. He terms this as ‘pyramid of gazes’ constituting new judicial process. In the UCL project, the camera was used to count different people and do surveillance of their movement. This concept of panopticon was again resurfaced by Michel Foucault in his 1975 book *Discipline and Punish*. His illustration was a way to suggest the civility and discipline in a society by its citizens. They have to follow certain instructions and orders in the name of civilised citizen. They are always conscious about their movement and act in certain fashion. He himself polices and is in fear of punishment or being judged. The cameras are the successor of the panopticon model where earlier watchtower used to guard the prison, and now through CCTV camera, surveillance is done. Foucault has emphasised on a mechanism that has permanent, exhaustive and omnipresent surveillance, that will remain invisible for others while maintaining its visibility ([Saifuddin, 2023](#)).

Unlike absolutist monarchical, authoritative gaze, that seeks for the control of entire mechanism, usually where all powers are vested in single person or group of people, in modern age, with the coming of AI all the

power are now classified under constant ‘unverifiable gaze’. Here the person is not the concerned player who will try doing surveillance on people but with the help of AI, s/he will just prompt the command and AI will automatically keep a record of all the movements and report with specific data. According to a BBC report published it says, “A camera system that uses AI and facial recognition intended to reveal states of emotion has been tested on Uyghurs in Xinjiang” ([Wakefield, 2021](#)). The article cited the claims made by a software engineer who has installed such software in the police stations in the province. China has invested heavily on the AI technology to keep surveillance in this province as it runs the risk of increasing extremism by the separatist group as they demand for their independence. It is a home to millions of detained Uyghurs Muslims under the ‘re-educational centre’ where they are treated under harsh rules and forced to study liberal religious texts.



(Wakefield, 2021)

Pie Chart-1

This pie chart shows the person's state of mind, red suggesting a negative or anxious state. But this data does not have any logical or conclusive

evidence to support this argument. He claimed it be “pre-judgemental without any credible evidence” ([Wakefield, 2021](#)). The China director of Human Rights Watch, Sophie Richardson terms it as shocking. She expresses her anguish “It's not just that people are being reduced to a pie chart, it's people who are in highly coercive circumstances, under enormous pressure, being understandably nervous and that's taken as an indicator of guilt, and I think, that's deeply problematic” ([Wakefield, 2021](#)). All the data are fed in the system and with the help of AI, it is used to generate algorithms that could flag suspicious activities. Such is the suspicion that people are monitored even for their back-door entry or filling gas in other's car. QR codes are placed on the doors so that they can easily figure who's inside and moving outside. In the name of safeguarding people, government or government backed companies monitor and store every minute details of a person. The phrase “Big Brother is watching you” stands perfectly aligned with this type of surveillance. George Orwell's dystopian fiction *Nineteen Eighty-Four*, published in 1949 talks about the use of power “for its own sake” over the inhabitants (Wikipedia [Big Brother Nineteen Eighty-Four, 2025](#)). Even Orwell wouldn't have thought about this kind of surveillance. Companies like Huawei, Hikvision are accused of sponsoring AI surveillance through their database on the behest of Chinese Government. One Person, One File was developed by the firm to store data of each person in their respective files. They could also store their political activities, relationships, ideology they support, which is a threat to life and freedom. AI detecting cameras by Hikvision and Dahua are accused of spying on Uyghur. Dahua though admitted to the fact that they have patents to monitor all 56 ethnicities in China and does not specific target the Uyghurs. The question arises from such claims that, why such data software and hardware are even needed. China is home to almost world's half of the CCTV cameras for surveillance, around 800 million. When the world is busy on developing smart cities that has pure water, good infrastructure, better jobs; Chinese smart cities such as Chongqing are equipped with AI environment. They will monitor each word we speak, every sign we make, each step we take, every person whom we meet; data of all this would be available to the authorities. With AI we have no safe heavens to hide. This could all sound like a Mission Impossible or James Bond Movie, but all this

is happening in real. Today it's China, nobody sure, tomorrow it may be Delhi, London.

In India Aadhar and PAN are seen and used in identification of individuals in different aspects such as opening of bank accounts, voting, ration list, etc. It has a unique identification number along with biometric details of each individual. In a recent debate that sparked the internet is the misuse of Artificial Intelligence, where Google's AI tool, Nano Banana, can be used to generate highly realistic looking identity cards. A Bangalore based techie, Harveen Singh Chadha, created this fake version of PAN and Aadhar through this app and shared it over Twitter. It highlighted the potential security risk that could lead to mass chaos and tampering with important information. Although it is too early to say that AI can bypass our servers but it qualifies to be a threat that can easily trespass physical verification done at hotel and airport.

In an ambitious run for the AI driven surveillance model for tracking up of suspected persons involved in crimes such as theft, murder, drug peddling, promoting terrorism, phishing, or monitoring of railway stations and bus stands for controlling of stampede like situation, we need AI, after all, it plays a crucial role in controlling these situations. But to the contrary, AI surveillance is used to track people who disagrees, speak against the government, institution, or powerful people. This institutionalisation of AI for silencing, suppressing voices is a serious concern that has to be checked upon. Article 21 of Indian constitution guarantees right to life and person liberty. In modern context, right to life and right to privacy, extend its context where it emphasises on the need to protect privacy digitally. India has framed a provision, Digital Personal Data Protection Act (DPDPA), 2023 where first legal framework has been laid for personal data protection. In this framework, entities that are responsible for storing, tracking, collecting digital data have defined obligations, that includes safeguarding and ensuring data accuracy. Data protection Board (DPB) has been set to track complaint and grievance. Non-compliance with the existing rules can lead to penalties and punishment. The penalties can go up to fine of 250 crores. Setting of Significant Data Fiduciaries (SDF) are designed for sensitive personal data that are processed. It is subjected to enhanced compliance requirement under DPDPA.

Instead of several regulations and safeguard measures, there are concerns and challenges looming over privacy and security. “Dragnet Surveillance” refers to the indiscriminate data collection which often extend beyond the suspects or criminals, potentially infringing citizen's privacy rights. Treating all as suspects, sends a warning for civil society and human right activists to demand privacy. Under DPDPA, government is exempted from its provisions under the guise of national security. Companies also take ‘mandatory consent’ when accessing its website, while showing that it respects privacy and collect personal data as per users wish. There are no AI specific guidelines leaving citizens vulnerable to potential misuse of AI technologies. While European Union (EU) has a comprehensive AI act, which acts at three levels. First is Unacceptable risk, that bans mass surveillance as in China by the government. Second is High-risk applications such as CV scanners, ID recognisers that runs identification, are subjected to specific legal requirements. Third is Minimal risk, which doesn’t fall in any category, are left unregulated. In India, we need to strengthen data laws that will have transparency and accountability, such as ‘right to explanation’ that will question the decisions made by the AI. And, if any fault is committed by the AI, who would be entitled for the punishment. Otherwise, if remain unchecked, it will be misused by the person in power. India stands at critical transition phase where it has to safeguard the privacy of its citizens by making robust laws, while maintaining security by AI driven models.

6 AI DRIVEN PHISHING AND CYBER FRAUD

Phishing is a practise of sending fraudulent communication that ‘appear’ to come from a legitimate and reputable source, usually through email and text messaging. The goal of attacker is to steal money, gain sensitive data and login information, or to install malware on victim's device (“What is Phishing,” n.d.). Phishing works by targeting the mind of victims, often looking legitimate and convincing. Sometimes ransomware or malware are downloaded onto the victim's computer. These phishing attacks looks so convincing as it exploits the vulnerability of human nature, including

tendency to trust others. They use different tactics like fear, greed, curiosity, sense of urgency, to compel recipient to open malicious links.

Agatha Christie's novel *The Moving Finger* (1943) featured a plot centred around anonymous foul, accusatory letters, disgustingly commenting upon the brother-sister relationship and terming them as lovers. This novel could be said to have been a precursor to modern day phishing and cyber bullying. These letters cause fear, disgust and suspicion in the mind of village people who talk sham about each other. "The letters are sent indiscriminately and serve the purpose of working off some frustration in the writer's mind. It's definitely pathological" ([Christie, 1943](#), 23). While Christie didn't use AI to solve crimes, her exploration of manipulative characters can draw parallel to understand how AI could be used for similar purposes. Same is with Phishing where the victim is caught and exploited, where malicious links try to deceive individual into revealing sensitive information and data.

From the regular phishing scam email and messages to the sophisticated AI driven, clean interface that looks almost real with cutting-edge technology to exploit human psychology with precision. The first phishing that occurred was in a chat room, many decades ago. Several types of phishing are committed such as ransomware, email account takeover (ATO), business email compromise (BEC), Whaling, Spear phishing etc. to name a few. Interesting fact stated by Tom Gillis, Cisco SVP and General Manager of Infrastructure and Security, noted that "80% of the ransomware attacks observed by Cisco started with a phishing email" ("What is Phishing," n.d.). AI has made it easier for attackers to compromise with user's data by sophisticated and targeted attack by writing correct and personalised emails or messages along with fraud links. This kind of personalised scam is called spear phishing. This is more dangerous than other forms as it is more tailored and specific, hence looks more legitimate. Cybercriminals collect personalised information and can use to withdraw, transfer money.

An infamous cyberattack, Whaling, where the famous CEO of Amazon, Jeff Bezos' iPhone was hit with a video of 4.4 mb on WhatsApp on May, 1, 2018 by Saudi Crown Prince Mohamed bin Salman (MBS) as they both recently exchanged their contacts. The phone began transmitting large amount of data without authorisation, according to Bezos' forensic team. It

is claimed that the massive cyber-attack was carried by NSO Group's Pegasus. This kind of technology doesn't come cheap. In past, Saudi Arabia has paid around \$55 million to the NSO group. The kind of phishing is meant to target journalists, princes of royal blood, high profile businessmen, high rank military and civil officials. Although there was no direct proof to validate this accusation. Saudi refuted such claims as baseless. Owner of Washington Post, Jeff Bezos was the prime target of the Saudi regime as its critic journalist Jamal Khashoggi was working there. Kevin Mitnick, one of the famous leading hackers called it "social engineering" to obtain key information like password. Stuxnet, a computer worm known for its malicious programming that seeps down the targeted supervisory control and data acquisition system. It was believed to target Iran's nuclear reactor site at Natanz, Isfahan province. The global electronics supply chain could be easily hit by malicious viruses, bugs and then supplied to the concerned parties. At least 32 people, including two children were killed, leaving thousands seriously injured after communication devices like pager and Walkie-talkie exploded across Lebanon in September 2024. These devices were used by Hezbollah operatives for their secure communication. Isreal is accused of plotting this conspiracy. Japanese brand Icom, responded to it by saying that "this handheld radio which was exported to the Middle East from 2004 to 2014 and has not been shipped since then" ([Tidy, 2024](#)). The production of this model was stopped 10 years ago. The battery provider of the device has also stopped its manufacturing. But some reports saying that the product is counterfeit. Prof. Simon Mahon, chair in International Relations at Lancaster University, told BBC: "we know that Isreal has a precedent of using technology to track its target, but called the scale of this attack unprecedented" ([Tidy, 2024](#)). This explosion was such tenacious that it shook entire Lebanon, exploding on several days. Phishing is the most common type of cyber-attack.

India has been the ranked second in phishing globally. AI based models are for precise targets unlike normal phishing attacks, which would send bulk messages to multiple targets. AI model has enhanced the success ratio of manipulating targets and exploiting them. Evading defense mechanism is easier for AI based scams. Zscaler Inc, a global leader in cloud security revealed that between January and December 2024, there were altogether two billion blocked phishing attempts. "Globally, phishing volumes

dropped by 20%, in part due to increased adoption of email authentication protocols. However, the nature of threats has evolved rapidly, with cybercriminals now using generative AI to deliver personalized lures, deepfake content, and malware-laced fake services on scale” ([ETCISO Desk, 2025](#)). India moved ahead of UK in terms of being most targeted cyber-attacks. Total 80 million phishing attacks were reported during 2024, accounting for 33% of global attacks. Phishing attacks are increasing on the social media sites like Facebook, Telegram, Twitter, and Steam. It is adding to malware attacks, gathering information and therefore carrying out “engineering attacks”. Attackers can use GenAI to increase attacks, generate fake websites and create deepfake texts, videos for social engineering. Prime targets are high stake government departments such as bank CEO's, director of defense wings, head of space and atomic departments, etc. cybercriminals are creating AI chatbots, AI assistants, that offers fake development of CV, graphic design, payment gateways, servers, hence collecting secured data and making them vulnerable.

AI technology is like walking on a double headed spear, where on one side it is the attackers which use AI in phishing to decoy and spread malware to extract vital information or money; on the other hand, agencies and private companies use the same AI to control the menace created by scammers. Along with active measures where security solution is deployed such as Cisco Umbrella for phishing protection and Cisco Secure Email Threat Defense to safeguard emails. The cyber industry should focus on increasing awareness about phishing. It is to train the workforce so that it can understand cybercrimes such as Phishing and malware, therefore protecting itself as well as company from throwing away of vital information. Security awareness and educational training (SAET) can reduce the number of cyber-attacks; however, it is near-zero chance to evade all cyber threats. As AI is improving the quality of messages that could not be easily caught by human intelligence, so differentiating a fabricated email and a genuine email would be very tough. Second, it will require robust knowledge and prompt awareness to protect themselves while maintaining genuinity in a communication channel. Third, it is difficult to fully trust SAET and AI driven programmes for countering attacks. Therefore, training with real-time threats and understanding the patterns used by scammers will reduce the impact when attacks do occur.

Cisco Umbrella DNS- layer security protects users from accessing malicious sites, links by blocking unsafe destinations, before any connection is ever made (“Block Phishing”, n.d.). Umbrella Secure Internet Gateway (SIG), further protects the users from phishing via unsafe URLs, non-standard ports and protocols, and more. “The proposed GenAI-SE (social engineering) framework can be used as a template to structure further research and serve as a baseline for empirical investigations into AI threats and countermeasures” ([Schmitt and Flechais, 2024](#)). It works as a threat template for professionals and data experts, providing support on both theoretical and empirical data to protect from AI-driven SE attacks.

7 FUTURE DIRECTIONS AND POLICY IMPLICATIONS FOR INDIA

While India's future looks promising in this digital revolution where Artificial Intelligence will make a crucial role in shaping the mind of our people, there needs certain regulations and checks to harness its true potential. There needs an AI governance which will look at ethical, policy making and awareness for countering the threats. Proactive measures such as acts like The Digital Personal Data Protection Act 2023, The Digital India Act (DIA) which will replace IT Act of 2000, The Telecommunication Act 2023, will regulate AI technologies within their ambit. Role of private sector is also crucial in regulating the AI as most of the data usage comes from there. Training people in smart usage of AI can only be performed when its in hybrid mode. India could adopt measures and laws from EU nations where its focus is on protecting people's rights while doing high-performance computing and making technological advancements. EU's Artificial Intelligence Act (AIA) which is a cornerstone in EU's regulatory system, can be beneficial for India's ambition in Artificial Intelligence.

8 CONCLUSION

With the high-end technological AI developments, it receives high volumes of digital crimes such as Phishing, deepfakes, AI surveillance, data manipulation, hacking, etc. It has deeply disturbed the mindset of common people while eroding the trust and security for effective internet use. People are sceptical about the increasing use of AI in day-to-day governance. This paper analyses the transition of crimes, when in twentieth century, crimes were committed such as murder, loot, acquiring of wealth through illegal means. But now people are concerned by the invisible, unsolicited crimes and criminals. This dangerous cocktail of AI crimes with the invisible players on the screen, hits below the belly. With the advance, more personalised attacks, it is nearly impossible to identify the threats and protect oneself in a ever evolving landscape of AI crimes. The “governmentality” approach in this cyber world possess the risk of being watched all the time, evading privacy. Strong cyber laws are the need of the hour with checks and balances, where government too is party to that. In the face of emerging technology, one need to be cautious about the changing patterns of crimes and need to develop robust countermeasures to the problem. Ethical, transparent and accountable regulations are needed for creating a ‘just’ online environment, in the ever-evolving era of AI technology.

REFERENCES

- Big Brother (Nineteen Eighty-Four). (2025, Aug 04). In Wikipedia. https://en.wikipedia.org/wiki/Big_Brother (Nineteen_Eighty-Four).↵
- Block Phishing Attacks with Cloud-based Cybersecurity. (n.d.). Cisco Umbrella. <https://umbrella.cisco.com/trends-threats/how-to-stop-and-block-phishing-attacks>.
- Can AI Solve Crime? (2025, April 17). BBC News. https://www.bbc.co.uk/learningenglish/features/6-minute-english_2025/250417.
- Christie, A. (1943). *The Moving Finger*. Harper Collins Publication.↵
- ETCISO Desk. (2025, July 22). *India ranks second globally and leads APJ in phishing attacks – Zscaler study*. ET CISO.

<https://ciso.economictimes.indiatimes.com/news/cybercrime-fraud/india-ranks-second-globally-and-leads-apj-in-phishing-attacks-zscaler-study/122834278>.[↵]

Foucault, M. (1994). *Power: Essential Works 1954-84*. Penguin Random House.[↵]

Foucault, M. (2020). *Discipline and Punish: The Birth of the Prison*. Penguin Classics.

Increasing Threats of Deepfake Identities. (n.d.). Homeland Security. https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf.[↵]

Interwar Britain. (2025, July 12). In Wikipedia. https://en.wikipedia.org/wiki/Interwar_Britain.

Laurent, L. (2020, Jan 24). *The Jeff Bezos Hack Shows Vulnerability Beyond Tech*. Bloomberg. <https://www.bloomberg.com/view/articles/2020-01-24/jeff-bezos-phone-hack-shows-vulnerability-beyond-tech>.

McMullan, T. (2015, July 23). *What Does the Panopticon Mean in the Age of Digital Surveillance?*. The Guardian <https://www.theguardian.com/technology/2015/jul/23/panopticon-digital-surveillance-jeremy-bentham>.[↵]

Saifuddin, S. (2023, January 07). *Reading Foucault: Panoptic Surveillance and Disciplinary Power*. The Companion. <https://thecompanion.in/reading-foucault-panoptic-surveillance-and-disciplinary-power>.[↵]

Sassoon, S. (1917). *Trench Duty*. In P. Jones (Ed.), *Modern British Poetry* (pp. 123–124). Macmillan.[↵]

Schmitt, M., Flechais, I. (2024). *Digital Deception: Generative Artificial Intelligence in Social Engineering and Phishing*. Artif Intell Rev 57, 324 Springer. <https://link.springer.com/article/10.1007/s10462-024-10973-2#data-availability>.[↵]

Sophie, L. (2024, May 21). *India's General Election is Being Impacted by Deepfakes*. Le Monde. https://www.lemonde.fr/en/pixels/article/2024/05/21/india-s-general-election-is-being-impacted-by-deepfakes_6672168_13.html#.

Tewari, D., and Sreejith, N. (2025, April 04). *Government of India Taking Measures to Tackle Deepfakes*. PIB.

<https://www.pib.gov.in/PressReleasePage.aspx?PRID=2119050>.

Tidy, J. (2024, Sep 20). *What We Know About the Hezbollah Device Explosions*. BBC News.

<https://www.bbc.com/news/articles/cz04m913m49o>.↵

Wakefield, J. (2021, May 26). *AI Emotion-Detection Software Tested on Uyghurs*. BBC News. <https://www.bbc.com/news/technology-57101248>.↵

Watson, C. (1971). *Snobbery with Violence: English Crime Stories and Their Audience*. Faber and Faber.↵

What is Phishing?. (n.d.). Cisco. <https://www.cisco.com/site/us/en/learn/topics/security/what-is-phishing.html>.

The Evolution of Privacy: Navigating Digital Surveillance

Nitesh Rai

University of Delhi

DOI: [10.1201/9781042018116-20](https://doi.org/10.1201/9781042018116-20)

ABSTRACT: We live in an increasingly interconnected world, where every single click, purchase, and interaction on social media necessarily generates a digital footprint. The idea of privacy faces complex challenges in such a scenario. Privacy, which initially began as a concern for inviolability of our physical space and sanctity of the home has transformed dramatically, entering into a new paradigm by the relentless march of new emerging technologies such as AI, ML, IoT, etc. With the emergence of internet, a new era began that marked a pivotal turning point, blurring the once clear lines between the public and private spheres. The Internet enabled exponential growth of digital platforms in the form of social media, search engines, and e-commerce websites, where we unleashed an unprecedented deluge of personal information voluntarily and often unknowingly. This enormous amount of personal data that we gave for free created a vast digital repository that gradually became a fertile ground for a new form of oversight: digital surveillance. As a result, the traditional concern for privacy that revolved around physical intrusion fundamentally shifted towards safeguarding control in an omnipresent virtual realm. In the context of privacy concerns, the objective of this chapter is to introduce the readers to the historical evolution of the idea of privacy, demonstrating its transformation from ancient philosophical ideals to its complex

manifestation in the contemporary digital age. This piece seeks to critically engage with some of the ethical issues and complex changes that we are exposed to with the advent of digital surveillance. The methodology that is used here is primarily qualitative content analysis and integrative literature review. Comparative analysis of case studies and policy review is employed for a comprehensive understanding. This chapter primarily relies upon the use of secondary sources in the form of articles, journals and news around several issues. This study uses cases from India, USA and China to underline the nexus between growth of technology and surveillance.

Keywords: Artificial Intelligence, Digital Surveillance, Privacy, Ethics, Machine Learning

1 METHODOLOGY

This study employs a qualitative research methodology that combines an integrative literature review with thematic content analysis. Secondary sources—including scholarly articles, legal texts, policy documents, and contemporary media—are systematically analysed to trace the historical evolution of privacy and its transformation under digital surveillance. A comparative case analysis of India, the United States, and China is employed to examine how technological growth intersects with ethical norms, governance frameworks, and surveillance practices across diverse political and regulatory contexts.

2 INTRODUCTION

The idea of privacy is not merely restricted to secrecy, rather, it should also be seen in a broader context where privacy enables an individual to control their personal information, autonomy to decide ‘when’, ‘how’, and to what extent the personal information is shared with other entities. In a way, privacy also embodies the character of individual rights, often treated as legal rights in modern discourse. This idea of privacy philosophically sees

privacy as an idea that is essentially about preserving human autonomy in an environment where information asymmetries create power imbalances ([*The Philosophy of Data Privacy: Why We're Building Programs That Miss the Point*, 2025](#)).

Digital Surveillance refers to the use of new emerging technologies such as Artificial Intelligence (AI), Machine Learning (ML), Internet of Things (IoT), etc., that monitor, track, and analyze individuals, groups, or entities, often without their knowledge ([Lee, 2020](#)). A whole range of activities is involved in digital surveillance, such as online data collection from several open platforms, social media, video monitoring, and biometric identification ([Holvast, 2009](#)). It is different from the traditional form of surveillance in the sense that it primarily focuses on the virtual world where much of our personal and professional life unfolds, unlike traditional surveillance, which primarily relied upon the physical environment ([Hawkes, 2024](#)).

The understanding of privacy has undergone significant changes with the passage of time, particularly concerning the locus of control. As society progressed, the innovations of the printing press, camera, and intrusive journalism posed a serious threat to someone's control over personal information. This resulted in a demand to treat privacy as a legal right, and privacy was interpreted as “the right to be let alone”. During World War I and World War II, new data collection tools (lie detectors, psychological testing, wiretaps) were discovered that had the potential to significantly undermine an individual's control over their information ([*Engaging Privacy and Information Technology in a Digital Age*, 2007](#); [Königs, 2022](#)). This journey from physical isolation to control over information is marked by the fact that the notion of privacy is a dynamic concept that is rapidly redefining itself in the age of technological innovations. New AI models have the potential to collect, share and visualize huge amounts of data across devices that are interconnected to each other. This, in effect, allows very little scope for us to control the flow of our personal information. This has further aggravated the power imbalance between individuals and entities that control this data flow. The open platforms (Meta, X (Twitter), Instagram, etc.), once cherished for the space of freedom and expression of opinion, have become a vast repository of personal information, blurring

the lines between online and offline lives and making personal data more accessible to governments and corporations.

3 EVOLUTION OF PRIVACY IN THE PRE-DIGITAL ERA: HISTORY AND PHILOSOPHY

In the discipline of Political Science, the earliest seeds of privacy can generally be traced to the Greek city-state, where a distinction was made between public (polis) and private (oikos). While the public space was mainly reserved for civic and intellectual engagements, household activities constituted the private domain ([Corsato, 2025](#)). The private space was typically reserved for women, slaves, and children and was accorded a lesser status as compared to public space activities. Citizenship as a title was granted to only those who engaged in civic and administrative activities of the city-state, the men. It is interesting to note here that this understanding of privacy stands in stark contrast to the modern idea of privacy that treats privacy as essential for individual freedom and autonomy. Romans viewed privacy through a different lens by emphasizing property and law. So, in contrast to the modern idea of personal privacy, where personal autonomy and confidentiality of thoughts are paramount, Romans, as observed by Esward Coke, saw privacy as *domus sua cuique est tutissimum refugium* (“each man's home is his safest refuge”) ([Coke et al., 1999](#); [Nwabueze and White, 2024](#)).

In the Middle Ages, although marked by kingship and dictatorship, religious institutions like churches, temples, and monasteries were places where people could go for seclusion, away from public scrutiny. In church, anyone could go and confess their crime without any worry of being identified. It should be acknowledged here that middle age was characterized by communal living and shared spaces; people resided in villages and monasteries, yet these religious places offered them a safe place to feel safe and away from the public gaze ([Cybulskie, n.d.](#)).

The major shift occurred in the 19th century when the emphasis on privacy inclined towards controlling personal information due to the rise of journalism exploitation of individual privacy for commercial interests. The

article, “The Right to Privacy,” published by Samuel D. Warren and Louis Brandeis in the 1890s in Harvard Law Review, was a direct response to the privacy issue exacerbated by instant photography and intrusive interviews that often capture and publicize private moments ([Warren and Brandeis, 1989](#)). Subsequently, what followed was new ways of personal data collection that included psychological testing and lie detectors, further intensifying the need for privacy. Alan Westin in 1967 redefined privacy as the right to decide “where, how, and to what extent” personal information is shared ([History of Privacy Timeline, n.d.](#); [Moskaleva, 2024](#)).

The notion of privacy has seen multiple interpretations because of changing social realities. The early understanding of privacy was restricted to privacy linked with private property, where private items were considered to belong to an individual's domain. This idea of privacy still resonates in many political discussions, and classical liberalism, in its exposition, has endorsed this idea. Privacy, as the “right to be left alone,” proposed by Warren and Brandeis, is still a cornerstone for legal interpretation. In the second half of the 20th century, several new interpretations of privacy emerged. Prosser, in his seminal work, divided the right to privacy into four torts: intrusion upon seclusion, public disclosure of embarrassing private facts, publicity placing an individual in a false light, and appropriation of name or likeness ([Solove and Richards, 2010](#)). However, if we consider these torts, they are insufficient and limited to protecting privacy in the physical realm. They are not adequate to address the extensive collection, visualization, use, and disclosure of personal information that is becoming a new normal in the information age. Different disciplines have defined privacy distinctly, and therefore, the notion of privacy has become multidimensional. For instance, in legal studies, the idea of privacy is distinct from what we believe privacy is in psychology, philosophy, or sociology. Privacy is defined as a feeling, emotion, limited access to the state, isolation, self-determination, power, and influence dynamics between individuals, groups, and institutions.

Despite this complex multidimensional approach to privacy, most philosophical works agree on at least five key aspects of privacy;

1. The idea of privacy revolves around individuals who claim it

2. The private sphere constitutes an individual to protect only a subset of their public information
3. An individual should have some degree of ownership over their public information
4. An individual should have the right to justify their claims concerning the protection of their public information, especially outside their private sphere.
5. The individual should be entitled to control the use of their public information, as privacy is also understood as the capability to control one's personal information to avoid risks of privacy loss ([Arora, 2019](#); [Benn, 1984](#); [Capurro, 2005](#); [Diggelmann and Cleis, 2014](#); [Privacy, 2023](#)).

Privacy is also understood as the ability of an individual to express themselves selectively to others while having control over the collection, use, and sharing of their personal information. However, if someone is intentionally making his personal information public, he, at the same time, also loses ownership and control over it concerning privacy.

This historical trajectory of privacy reveals an enduring tension between the public and private spheres and between social utility and individual autonomy. From the early distinction between *oikos* (private sphere) and *polis* (public sphere) to the modern challenge of balancing the utility of data collection with individual autonomy, this tension has remained constant. Thus, what we see here essentially is that the fundamental nature of what is being protected has evolved from physical isolation to control over information. However, amidst all this, the core tension between individual rights (seclusion, autonomy, control over personal information) and collective good (public order, innovation, national security, commercial gain) persists.

This reinforces the proposition that privacy is not a fixed concept; it is rather an ongoing negotiation within societal values and technological prowess. Legal response to this problem has inherent issues, that is, new emerging technologies are constantly changing and learning, thereby posing new challenges very frequently. Our legal system often does not correspond to the pace at which this constant change is occurring. In a time when our personal data is mined extensively in different forms, even the existing torts

like “right to be let alone” are difficult to implement in a true sense. A flexible legal system that is open for new updates in accordance with changes in technology is the need of the hour; otherwise, the very foundational definition of privacy that is “right to be let alone” would be restricted to philosophy only. This would severely undermine the rights of everyone, especially those who are already marginalized.

4 THE ADVENT OF DIGITAL SURVEILLANCE

Surveillance in its initial stage primarily began with electronic surveillance when telecommunication emerged as a prominent tool to gather long-distance information more effectively. Use of telephone, closed circuit television (CCTV), telegraph and camera were the tools that played a very prominent role, thereby allowing law enforcement agencies to widen their surveillance. CCTV cameras enabled police to watch, record, and track a large number of people in public spaces. Telematics (merger of computing and telecommunication) has helped in service providing, for instance, tele-medicine and tele-education, but it also poses a significant risk of distribution and misuse of personal data. The period of the 1970s thus saw the creation of privacy-enhancing technologies (PETs) being developed by researchers and scientists. Andrew Yao (multi-party computation) and David Chaum (blind signature) were some of the notable contributions in the direction of PETs development. This period was also marked by protest, resistance, and privacy-focused movements ([Chaum, 1983](#); [Evans et al., 2018](#)).

The cypherpunk movement that began in the late 1980s sought to ensure the deployment of strong encryption and privacy-enhancing technologies (PETs) to protect the privacy of individuals; at the same time, it also aimed at resisting state or corporate surveillance ([Ramiro and de Queiroz, 2022](#)). It was a monumental shift in the use of encryption technology because, thus far, the use of encryption was the prerogative of the state only. Pretty Good Privacy (PGP), a software that allowed users to protect their personal messages, was a landmark moment, invented by Phil Zimmermann in 1991.

Very soon, it established itself as an embodiment of privacy by allowing people to encrypt their personal messages despite the fact that the US government was not in favor of this ([Russakovsky *et al.*, 2004](#)). Digital surveillance, however, continued with more advancements and sophistication with the passage of time. Now, it involves tracking cookies, IP addresses, and the volume of internet traffic in order to monitor and surveil people worldwide. It has effectively rendered most of our online activity susceptible to being traced and monitored which is a serious breach of the notion of privacy.

The contemporary landscape of digital surveillance is characterized by massive employment of AI, IoT, quantum computing, advanced biometrics, brain computer interface and cloud computing. What makes these new technologies different from their predecessors is their unique capabilities to collect, store, and visualize personal data at an unprecedented scale which was unimaginable before few years ago. AI has truly brought revolution in the field of electronic surveillance by assisting AI models to detect patterns, predict, visualize, and analysis of our behavior from massive data that it accumulates from CCTV footage, Facial recognition, sensors and social media platforms. In the following section, through case studies and illustrations, the application, violation of privacy and other major significant ethical issues are raised with critical engagement.

Facial Recognition Technology (FRT) use nowadays is widespread for various reasons. It is basically a biometric identification system that helps to match images and facial features captured through CCTV cameras with a huge repository of images stored from multiple platforms, such as railways, airports, social media, street images, etc.

It is primarily, but not limited to, used by enforcement agencies for the prevention of crime, suspect identification, and tracking of suspected individuals. On its face, it may seem the use of a machine for such a novel cause; however, the other side of FRT is suppression of dissent, illegal tracking and monitoring of individuals, creating fear in their minds, thus overall subversion of human dignity and autonomy that is detrimental to any democratic setup. There are several cases of misuse of FRT recorded across the countries where civil rights are denied by the identification of protestors, harassment, and victimization are carried out by the state using FRTs. Its inherent prowess for mass surveillance, identification, and data

collection without even taking any explicit consent from the individual makes it dangerous and against the freedom of the individual.

In 2019, during the Hong Kong protests, the use of FRT allowed enforcement agencies to track individuals in real time. China used FRTs to utilize enormous amounts of video and images to identify protestors, which helped it in massive crackdowns on protestors ([Zalnieriute, 2024](#)). This created a fear psychosis in the minds of the protestors, which in turn deferred their active participation in the demonstration. Although in response to this, protestors resorted to new methods as a decoy to escape FRTs by wearing masks, goggles, using handheld laser pointers, aiming cameras to blur the recordings, and destroying those lampposts that protestors suspected of having surveillance cameras ([Rhodes, 2025](#)).

In August 2019, demonstrators toppled these lampposts in Kowloon, turning them into symbolic targets in their battle against surveillance. Protesters also creatively improvised using common items like aluminum foil, reflective materials, and thermal blankets to shield themselves from heat signatures and reflect images to cameras. Protest is, by no stretch of imagination can be understood alien to a healthy democracy. But when state machineries resort to such subversive tools to track and penalize protestors, this undermines the very basic ethos of democracy, freedom of speech, and assembly. The chilling effect of such incidents often makes citizens conforming to state laws that they do not even like, but the fear of being identified and subsequently penalized forces them to do so ([Fussell, 2019](#); [Hu, 2019](#); [Rhodes, 2025](#)). This would certainly not be appreciated in any democratic order. While it is true that after facing criticism from all quarters, China has now made it illegal to mandate facial recognition for availing services, such regulation primarily focuses on consumer protection and consent in a commercial sense; it is not against broad state surveillance ([Interesse, 2025](#)).

In India, the National Crime Record Bureau (NCRB) under the Ministry of Home Affairs is also establishing National Automated Facial Recognition System (AFRS) that aims to create a centralized national database of images accessible to law enforcement agencies for the identification of individuals, criminals, missing persons, and unidentified bodies (*Automated Facial Recognition System Will Facilitate Better Identification of Criminals, Unidentified Dead Bodies and Missing/Found*

Children and Persons: Shri G. Kishan Reddy, 2020). However, the apprehension that AFRS may perpetuate illegal state-sponsored mass surveillance cannot be ruled out, which can have wider ramifications and chilling effects on the right to privacy that is now recognized as a fundamental right via Supreme Court landmark judgement *K.S. Puttaswamy v. Union of India* in 2017. These apprehensions are not merely speculative because even before this, AFRS could be operationalized, we have an instance where Delhi Police used a facial recognition system to trace a missing child in the case *Sadhan Haldar v. The State NCT of Delhi* [WP(Crl) 1560/2017]. As per the Indian Express report, the software was sourced from M/s Innefu Labs Private Limited. However, in reality, this did not yield much success, with at least three orders of the Hon'ble High Court recording its dissatisfaction with FRTs. The court found it “unacceptable” that facial recognition had not helped to trace any of the 5,000 children missing from the city in three years. From the Madurai city police of Tamil Nādu, phone numbers and photos were leaked from its facial recognition database ([Chaudhary, 2019](#)).

In the absence of any clear court ruling regarding under which circumstances FRTs can be used and what the SOPs are for their use, their use still remains in the grey. Because of its high inaccuracy along with misuse by enforcement agencies, it may cause a biased approach and coercive actions ([Jain, 2019](#)). It is interesting to note here that Delhi Police acquired this software in 2018 as a tool to identify lost and found boys and girls by matching photos, but it was used on December 22, 2019, in a political rally of the Prime Minister to screen the crowd ([Mazoomdaar, 2019](#)). In the year 2020, Home Minister Amit Shah told Lok Sabha that authorities used facial recognition from Voter ID cards, driving licenses, and other databases to identify that 300 people involved in the Delhi riots ([Mehrotra, 2020](#)).

India is planning to set up one of the world's largest facial recognition systems in Hyderabad. The city police, in its announcement, revealed that the number of CCTVs in Hyderabad will double to 1 million by the end of 2020, which means that, on average, there will be one camera for every 10 citizens in the city. The police have presented this entire episode as it is better to be safe than sorry. A digital rights activist from the city admitted that in a state like Telangana, where protests are not allowed, “the police

tend to record everyone during protests and then come after them after the political momentum fades away”, thus making the apprehension true ([Bansal, 2020](#)).

The underlying problem, apart from misuse, is also the absence of any data protection or privacy law. A person whose company was responsible for developing facial recognition for eight local police forces acknowledges that India does not have the same quality of camera as China, thus rendering it hard to meet the goal of identifying anyone with accuracy from a database. Hence, in the absence of a comprehensive privacy protection law combined with proposed nationwide deployment of FRTs and considering the high inaccuracy rate, this all makes a perfect pathway for possible misuse of human rights. This is not just about individual violations of privacy; it also represents the potential for the state to create a panopticon that inevitably will target, monitor, surveil, and possibly criminalize marginalized populations. Such a scenario carries with it a profound chilling effect on civic engagement.

In the USA, people are coming together in the growing movement to ban police use of facial recognition because this technology is simply too dangerous in the hands of police. According to a news report in the Washington Post, two men who were so far unknown to the public, namely Galtin and Jason Vernau, were wrongfully arrested in St. Louis and Miami after being misidentified by facial recognition ([MacMillan et al., 2025](#)). When it comes to the Black community in the US, face recognition specifically has a long history of having a lower rate of accuracy because of darker complexions. In a study conducted in 2018, Joy Buolamwini and Timnit Gebru demonstrated that facial recognition technology has a significantly higher rate of error, especially for dark skinned women (34.7%) as compared to light-skinned men (0.8%) ([Fergus, 2024](#)). Black people are being victims of this racial bias because of the risk scoring system. This software, Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is used across the country for predicting future criminals from pools of personal data, and previous criminal track records are grossly biased against black people. COMPAS is a software that is developed by a private agency Northpointe.Inc.. and its main purpose is to predict.

The likelihood of a crime an individual would commit in future. It also supervises and recommends what kind of punishment one should get based on risk assessment result by the software. COMPAS uses the past data of an individual, research before giving any prediction. The company who is responsible for making this algorithm claims that men, women and juveniles receive different risk assessment score but how it is done is kept secret. Although this risk assessment application is not the sole criterion to pronounce judgment, judges do infer inputs from this. If the machine at the very first instance scores someone in poor light, the accused has every chance of being deprived of the assumption of innocence, which is the basic requirement of every fair trial. Judges are factoring the risk assessment score of an accused to decide on the terms of their bail, sentence, rehabilitation, and detention ([Rai, 2022](#)).

The threat is looming large on those people who already have a history of some crimes, as the machine would be using those past records to predict their future involvement in crime. An actual demonstration would be better to elucidate my points here. Below is a picture in which Angwin and others have discussed a case where two people with different criminal acts were imprisoned. A 41-year-old Vernon Parter, who is a white guy, who had already committed two armed robberies and one attempted armed robbery, was booked in a case of shoplifting. Brisha Borden, who is an 18-year-old girl, was booked for petty theft worth 80\$, that too she did unintentionally. When it comes to risk assessment, the machine risked Brisha more than Vernon. Brisha scored eight and was put in the high-risk category; on the contrary, Vernon scored three and was placed in the low-risk category by the software. What is more interesting to note here is that Vernon was again caught in a subsequent grand theft case, but no such act was done or even attempted by Brisha Borden in the future ([Angwin et al., 2016](#)).



Figure 1. Petty theft arrests.

The persistent occurrence of wrongful arrests disproportionately implicating Black people in the United States illustrates the idea that FRTs not only sustain bias but also amplify the racial bias already embedded in the criminal justice system. It transforms biased historical discrimination into a new technological form of systematic discrimination. Another concern that is directly related to digital surveillance and violation of privacy is behavioral analysis and predictive policing performed by AI-assisted machines. The Internet of Things (IoT) refers to all such devices, including sensors, cameras to our wearable watches and smart home appliances, that are connected over internet and allow for a deeper level of monitoring of our environment and activities of ours'. Without any iota of doubt, IoT has enabled us to enjoy luxury and utilize our time in some other productive works, it has also raised the threat bar because of seamless transfer of our data across the interconnected devices. A large concern is

that the devices are connected to the internet and may be hacked, which could then result in numerous privacy breaches and access to sensitive personal data.

5 DEMOCRATIC VS. AUTHORITARIAN LOGICS OF SURVEILLANCE

Persistent use of AI in the broader domain has sparked another discussion regarding the nature of AI models in the future. China, for example, has often been described as a closed country that breeds an authoritarian regime and provides minimal opportunity to speak against its leadership and their actions. ([Alvin Y., 2019](#); [Zhang and Mora, 2023](#)). Any AI system will necessarily be anti-democratic and authoritarian in nature, which is a natural corollary of this wider belief. It is assumed that a country like China will use machines to further surveil, control its population and crush dissents, if any. This supposition has not developed in silos. We have discussed in the preceding section how China used AI-assisted technologies to control and crush the Hong Kong protests. Its use of an AI model for social profiling has raised pressing concerns for those who care for human rights and dignity ([Chen, 2024](#); [Zhao and Liu, 2025](#)). But then we should also raise the question and see the status of the AI model, the purpose of deployment in widely believed democratic countries as well. Even if we accept the proposition that USA and India are more democratic than China, has this assumption really translated into reality when it comes to these countries using AI for different purposes? In the same preceding section, we have discussed cases of India and USA with respect to FRTs and COMPAS and by no stretch of imagination can we align the outcome of these technologies with standard democratic principles.

There are many other incidents in countries which are often ranked very high on the democracy index. Germany that has secured 1st place in annual global rankings of country's democratic performance ([Country Rankings, n.d.](#)), has reportedly accepted the EU AI Act that will propose significant exemptions for securities agencies. This is interpreted as detrimental to people's rights. Germany was critical of remote biometric identification

(RBI) initially and stressed that “biometric identification through AI in public spaces must be ruled out by EU legislation.” However, within a few months, it changed its stance and stressed that only ‘real-time’ RBI should be forbidden, and only law enforcement authorities or their subcontractors would be subject to these limitations. This effectively gave free hand to private entities and public authorities other than law enforcement to use RBI. The ban on RBI also does not apply if any analysis is done on stored data because only real time data is prohibited from access ([Aszódi and Spielkamp, n.d.](#)). Further, a number of exemptions are granted to law enforcement agencies to use RBI on the grounds of prevention of threats to health, public safety, law and order. What does this show? Irrespective of regimes labelled as democratic or non/less democratic, the tendency of states to use AI model to enhance their power is a common trend. States initially justify the amplification and intrusion of AI models to surveil and control the population on some extraordinary grounds, like health and public safety. However, such practices will soon be normalized in our day-to-day lives. As more countries are gaining knowledge and mirroring the development and deployment rate of advanced countries, it would be crucial to watch the direction and nature of AI adoption.

6 THE INTERSECTION OF PRIVACY AND DISCRIMINATION

Generally, privacy and bias are treated as two different concepts, and we rarely combine them. In this section, an attempt is made to see privacy and bias together, especially in the light of AI-generated outcomes. For AI to work, data needs to be taken from various sites. In this whole process, our personal data, such as gender, political ideology, caste, color, and financial status, may be taken without our knowledge. Now imagine a situation in which an AI system, because of bias in its algorithms, is reflecting prejudices against an Individual.

Let us assume the machine is biased against a financially weak person in its prediction. By the way, it is happening now. ‘Thick file’ and ‘Thin file’ are associated with this. Here, what is interesting to observe is the correlation between privacy and bias. It could exhibit a biased prediction and could corner out all those who were financially weak because of the data it had. So, in the first place, as an individual, you may not be aware that your financial profiling is being done based on your financial activities. The data breach is established here. Eventually, this data is being used to discriminate against you. To isolate you. Take an example of an AI that an employer may be using to examine job applications. If there is systemic discrimination against women or marginalized racial or ethnic groups, it is possible that the data that is being captured about a candidate's gender or ethnicity may be used against them to unfairly eliminate them from the evaluation in an unjust manner. This has detrimental effects on the individual candidate and contributes to the continuation of structural disparities in the workforce. We are all aware that because of automation in industries by smart machines, there is a massive displacement of workers, creating unemployment on a large scale. If we look closely at this, we see that unemployment is linked to the issue of privacy, just like bias is related to privacy.

Now that we have seen through several case studies that new technologies are challenging the societal structure and new concerns for privacy and discrimination are on the rise, the following section will

critically analyze policy discussion and address these challenges to some extent.

Among many other concerns, a primary concern is the huge amount of data that is often captured and stored by private companies; the possibility of sharing this data across borders further aggravates this problem ([Rosen, n.d.](#)). Besides law enforcement agencies that often illegally sneak into the database for purposes other than defined by the court of law, corporate players also use this personal data for commercial purposes like advertisement, customization, and selling. Despite all these concerns, we observe a slow response from legal and legislative bodies to adapt to such changes and frame regulations accordingly. Apart from highlighting vulnerabilities that lie at the user's end as far as control over the flow of personal data is concerned, this data-driven economy can exacerbate existing inequalities in society. For example, rich people may afford to shield themselves from intensive digital surveillance by using privacy-enhancing tools and technologies; however, not everyone can afford it.

Thus, your economic status gives you options and greater freedom to shield your privacy as compared to those who cannot afford it. For any civilized society, privacy should be a basic human right. It should have nothing to do with your economic status and background. Another concern that policymakers should be mindful of is the normalization of surveillance and privacy violations.

Digital surveillance is often presented as a tool that can be used for both beneficial and harmful purposes. For example, surveillance can be used to improve public safety, but it can also be employed by the state or security agencies to suppress dissent or protest against the establishment.

Electronic surveillance is often justified in the name of national defense, security, law and order and prevention of crime. Even a stringent EU act to regulate AI provides for exemption on health and law and order grounds. The reasoning for this, as argued by security agencies, is that they need to use this smart application to detect the pattern and prevent a crime before it can actually take place. In doing so, what we have commonly observed is right to privacy and breach of personal space are being undermined and outweighed by the anticipation to protect public safety by the security department. While it is a well-accepted phenomenon that there could be

security concerns that require some forms of surveillance, the point is that these justifications cannot keep extending without any critical assessment.

Further, when we ask this question, how can society balance the immense advantages being offered by AI machines that improve our overall societal well-being with the concern of protecting our individual privacy at the same time? The argument is that we should not see this issue as either/or; instead, we should see this whole issue as a matter of finding frameworks for responsible engagements that can protect our privacy at the same time, allowing for the beneficial use of data. This should not be seen as a zero-sum game; instead, a more responsible engagement can help here. A multifaceted approach is needed to regulate surveillance, which includes stringent legal protection for the privacy of individuals, greater transparency and accountability of both state and corporate, and engagement of the public in debates about surveillance and technology. The vital role of international Corporations cannot be ignored in developing global norms and standards because challenges posed by surveillance are increasingly becoming global in nature. We have to understand that surveillance is not a static phenomenon; rather, it is a dynamic and evolving process, and will continue to evolve; therefore, our response should also be dynamic and evolving in nature. Tech corporations should be encouraged to adopt popular data safety methods in the form of Privacy by Design (PbD), regular risk assessment, transparency in data processing, and a clear, unambiguous data retention policy.

7 STRATEGIES TO NAVIGATE DIGITAL SURVEILLANCE

The following list of recommendations is not exhaustive; however, they do offer a general understanding and awareness to users and corporations about methods that can be used to protect privacy and avert any unwanted consequences with the use of AI, ML, IoT, etc...

1. Emphasis should be given to Privacy-enhancing technologies (PETs), as they are meant to implement data protection principles, minimize misuse of personal data, and ensure that technologies are in accordance

with the local law of the land. PETs include encryption, which is crucial for privacy, and act as doorkeepers by preventing anyone from seeing our personal data when transmitted through networks. Anonymization techniques and data masking are important components of PETs. In anonymization, we hide or remove personally identifiable information (PII) from the database while ensuring data utility for testing ([Han and Lu, 2024](#)).

2. Differential privacy is another method that adds some randomness or noise into query responses from a dataset, thereby making it next to impossible to pinpoint individual data points. Thus, without disturbing data analysis, it protects individual personal information by making it statistically hard to pinpoint.
3. Homomorphic encryption (HE) is nowadays used as a potential cryptographic technique that enables computations to be performed directly on encrypted data, and there is no need for decryption anywhere. This is game-changing for secure cloud computation, secure AI/ML training on sensitive data, and compliance with regulations ([Ajala et al., 2024](#)).
4. At the individual level, the use of secure networks and virtual private networks (VPNs) creates a secure and encrypted connection between the device and the server, thus potentially undermining the risk of hacking and data breaches.
5. The limited use of social media, where we litter enormous amounts of unnecessary information for public consumption, should also be minimized judiciously. Use of end-to-end encryption tools should be encouraged for all possible communication and transmission of personal data. Many internet browsers offer private engine search, do not track, clear private data, and anti-tracking features, that can help to protect our online behavior and personal data.
6. Data minimization should also be practiced to its optimum level because it allows the collection and retention of only the personal information that is directly relevant and essential for a specific work. The data is then removed from the repository when its purpose is achieved ([Senarath and Arachchilage, 2019](#)).

8 GLOBAL GOVERNANCE OF DIGITAL SURVEILLANCE

Realizing that this is slowly turning into a global problem, several countries have taken noble initiatives to ensure a fair degree of balance between innovation and risk. Letting AI not be allowed untrammelled flow in the absence of any regulation is a common discussion point on which we all agree today. Consequently, different regulations aiming to make AI models ethical, accountable and secure for users emerged. The EU AI Act is considered to be a strenuous effort in the broader direction of safe AI for users. It categorizes AI machines into different categories depending upon their consequences for the public and accordingly puts forth strict regulations to adhere to. In case of non-compliance, severe financial penalties are clearly stipulated in the act. In general, this act also bans live facial recognition in public spaces, emotion recognition and biometric categorization. Exemptions are provided for law enforcement for serious crime, and medical/safety purposes only ([Artificial Intelligence Act: MEPs Adopt Landmark Law, 2024](#)). OECD AI guidelines for safe AI mandates for the creation of robust, safe and secure AI, including provisions for override, repair and decommissioning any AI system if it poses any serious threat to human safety for some reason, including undesired behaviors (*AI Principles*, n.d.). Countries like the USA and China have also pioneered in making multi sectoral and centralized regulations for AI models, respectively. While it should be appreciated that responsible countries and organizations are thinking in this direction, we are witnessing the reluctance of new emerging economies when it comes to either framing regulations or adopting the guidelines for some obvious reasons. One predominant reason is that any strict enforcement of AI laws may stifle their newly growing AI ecosystem set-up. As advanced countries have historically denied transfer of technology to the needy once, new emerging markets have either put such plans in abeyance or performing perfunctory action in trepidation, afraid of losing a potential opportunity. This fear is a predisposing factor for the slow adoption of AI regulations in many countries. For this, a basic set of laws should be prepared with consultation with all stakeholders, here all UN countries. This may probably be the first step to thrust further serious

attempts to make AI safe for public consumption by countries who are sluggish so far.

The regulatory mechanism has to ensure that concerns from diverse stakeholders are included in any guideline to regulate AI. The present-day reality shows that private entities like Meta, Google, and tech giants are as much involved as any state government is, and therefore, they too should be nudged to think above profit-making tendencies only. Shoshana Zuboff has already flagged the modus operandi of Google, where she highlights how Google has treated the personal data of its users as another site for profit-making with complete disregard for privacy concerns ([Zuboff, 2019](#)). At the same time, even the most stringent law of its kind, the EU AI Act, alone would not be sufficient because of the cross-border flow of data. For any effective global guidelines for AI should incorporate views from individuals, state and non-state entities (Meta, X, Google, etc.), global institutions like the IMF, World Bank, UN agencies dedicated to protecting human rights, and especially voices from the global south. We may not achieve complete success in our first attempt, but if we are consistent with our effort in this direction, we can certainly avert many ethical issues arising out of AI use at various places.

9 CONCLUDING REMARKS

The trajectory of privacy from physical seclusion to the present struggle for control over personal information highlights a fundamental truth: the notion of privacy is continually redefined by technological advancement as well as the evolution of social norms. This chapter has mapped the evolution of privacy by showing how the understanding of privacy from “right to be let alone” has transformed into a complex challenge of struggling to manage vast digital footprints. The proliferation of AI, ML, cloud computing, and other advanced technologies has ushered in an era of digital surveillance, alteration of the public and private spheres, and amplification of power asymmetries between individuals and entities responsible for data collection. The chapter has underscored profound ethical challenges through the illustration of several case studies, ranging from the suppression

of dissent via FRTs in Hong Kong, bias in algorithms of predictive policing systems like COMPAS in the USA, to the grey area of misuse of these technologies in India in the absence of any unambiguous regulation. These technologies not only threaten the dignity and autonomy of an individual, but they can also amplify and reinforce existing social inequalities in our society, if not ethically audited.

The normalization of surveillance, often disguised as security, has a chilling effect on civic engagement and poses a threat to democratic freedoms. The already significant gap between those who have the means to utilize privacy-enhancing tools and those who do not is becoming even starker, meaning that privacy, an inherent human right, is becoming a privilege. The solution to this complex challenge can never be a zero-sum game where demand for privacy outweighs innovation or vice versa. It rather demands a multifaceted, proactive approach where innovation is fine-tuned with a robust privacy framework, strict legal protection, and an adaptive legal and legislative process. Transparency, public accountability, and awareness of digital literacy will help strengthen our resolve to make this world safer and more peaceful. The onus of this is on policymakers, technologists, tech giants, and, very importantly, citizens to ensure that technological advancements serve the best interests of humanity by ensuring the right to privacy is harmonized with innovation and privacy remains an unassailable cornerstone of a free and just society.

REFERENCES

- AI principles. (n.d.). OECD. Retrieved January 11, 2026, from <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>
- Ajala, O.A., Arinze, C.A., Chrisanctus Ofodile, O., Chinazo Okoye, C., and Daraojimba, O.D. (2024). Reviewing advancements in privacy-enhancing technologies for big data analytics in an era of increased surveillance. *World Journal of Advanced Engineering Technology and Sciences*, 2024(01), 294–300. <https://doi.org/10.30574/wjaets.2024.11.1.0060>

- Alvin Y. (2019). The rise of authoritarianism in China in the early 21st century. *International Review of Modern Sociology*, 45(1 Special Issue), 53–75. <https://researchportal.hkust.edu.hk/en/publications/the-rise-of-authoritarianism-in-china-in-the-early-21supstsup-cen/>
- Angwin, J., Larson, J., Mattu, S., Kirchner, L., and ProPublica. (2016, May 23). *Machine Bias*. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Arora, P. (2019). General data protection regulation—a global standard? Privacy futures, digital activism, and surveillance cultures in the global south. *Surveillance and Society*, 17(5), 717–725. <https://doi.org/10.24908/SS.V17I5.13307>
- Artificial Intelligence Act: MEPs adopt landmark law. (2024). *European Parliament*. <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>
- Aszódi, N., and Spielkamp, M. (n.d.). *How the German Government Decided not to Protect People Against the Risks of AI*. AlgorithmWatch. Retrieved January 11, 2026, from <https://algorithmwatch.org/en/german-government-risks-of-ai/>
- Automated Facial Recognition System will Facilitate Better Identification of Criminals, Unidentified Dead Bodies and Missing/Found Children and Persons: Shri G. Kishan Reddy. (2020, March 4). Press Information Bureau. <https://www.pib.gov.in/PressReleasePage.aspx?PRID=1605148&reg=3&lang=2>
- Bansal, V. (2020, November 10). *The Hyderabad model of CCTV surveillance*. Mint. <https://www.livemint.com/news/india/the-hyderabad-model-of-cctv-surveillance-11604926158442.html>
- Benn, S.I. (1984). Privacy, freedom, and respect for persons. *Philosophical Dimensions of Privacy*, 223–244. <https://doi.org/10.1017/CBO9780511625138.009>
- Capurro, R. (2005). Privacy. An intercultural perspective. *Ethics and Information Technology*, 7(1), 37–47. <https://doi.org/10.1007/S10676-005-4407-4>
- Chaudhary, A. (2019, September 20). *Facial recognition Plan: India is planning a huge China-style facial recognition program*. *The Economic*

Times. <https://economictimes.indiatimes.com/news/politics-and-nation/india-is-planning-a-huge-china-style-facial-recognition-program/articleshow/71213132.cms?from=mdr>

Chaum, D. (1983). Blind signatures for untraceable payments. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 199–203. https://doi.org/10.1007/978-1-4757-0602-4_18

Chen, Y. (2024). Disregarding Blameworthiness, prioritizing deterrence: China's Social credit-based punishment and the erosion of individual autonomy. *The China Review*, 24(3), 139–177.

Coke, Edward., Butler, Charles., Hargrave, Francis., and Littleton, Thomas. (1999). The first part of the institutes of the laws of England, or, A commentary upon Littleton : Not the name of the author only, but of the law itself. https://books.google.com/books/about/The_First_Part_of_the_Institutes_of_the.html?id=EVUkwWB052cC

Corsato, B. (2025, October 30). *The origins of privacy in Europe*. Usercentrics Magazine. <https://usercentrics.com/magazine/articles/geography-of-trust-origins-of-privacy-in-europe/>

Country Rankings. (n.d.). Global State of Democracy Initiative. Retrieved January 11, 2026, from <https://www.idea.int/gsod/2024/countries/>

Cybulskie, D. (n.d.). *Privacy in the Middle Ages*. Medievalists.Net. Retrieved January 11, 2026, from <https://www.medievalists.net/2023/09/privacy-middle-ages/>

Diggelmann, O., and Cleis, M.N. (2014). How the right to privacy became a human right. *Human Rights Law Review*, 14(3), 441–458. <https://doi.org/10.1093/HRLR/NGU014>

Engaging Privacy and Information Technology in a Digital Age. (2007). National Academies Press. <https://doi.org/10.17226/11896>

Evans, D., Kolesnikov, V., and Rosulek, M. (2018). A pragmatic introduction to secure multi-party computation. *Foundations and Trends® in Privacy and Security*, 2(2–3), 70–246. <https://doi.org/10.1561/33000000019>

Fergus, R. (2024, February 29). *Biased Technology: The Automated Discrimination of Facial Recognition*. ACLU Minnesota.

<https://www.aclu-mn.org/news/biased-technology-automated-discrimination-facial-recognition/>

Fussell, S. (2019, August 30). *Why Hong Kong Protesters Are Cutting Down Lampposts* - The Atlantic. The Atlantic.

<https://www.theatlantic.com/technology/archive/2019/08/why-hong-kong-protesters-are-cutting-down-lampposts/597145/>

Han, Y., and Lu, X. (2024). Is data anonymization an effective way to protect privacy or not. *International Journal of Computer Science and Information Technology*, 4(3), 152–156.

<https://doi.org/10.62051/ijcsit.v4n3.15>

Hawkes, P. (2024, December 9). *Digital Surveillance Explained*. Research Associates.

<https://researchassociates.com/digital-surveillance-explained/>

History of Privacy Timeline. (n.d.). *Information and Technology Services Safe Computing*, University of Michigan. Retrieved January 11, 2026, from

<https://safecomputing.umich.edu/protect-privacy/history-of-privacy-timeline>

Holvast, J. (2009). History of privacy. *IFIP Advances in Information and Communication Technology*, 298, 13–42. https://doi.org/10.1007/978-3-642-03315-5_2

Hu, C. (2019, September). *What Hong Kong's masked protesters fear*. CNN World. <https://edition.cnn.com/2019/09/09/asia/smart-lamp-hong-kong-hnk-intl>

Interesse, G. (2025, March 31). *China's Facial Recognition Regulations: Key Business Takeaways*. China Briefing. <https://www.china-briefing.com/news/china-facial-recognition-regulations-2025/>

Jain, A. (2019, December 29). *The Delhi Police must stop its facial recognition system*. Panoptic Tracker. <https://panoptic.in/case-study/the-delhi-police-must-stop-its-facial-recognition-system>

Königs, P. (2022). Government surveillance, privacy, and legitimacy. *Philosophy and Technology*, 35(1), 8. <https://doi.org/10.1007/s13347-022-00503-9>

Lee, R.S.T. (2020). A brief journey of human intelligence. In *Artificial Intelligence in Daily Life* (pp. 3–18). Springer Singapore. https://doi.org/10.1007/978-981-15-7695-9_1

- MacMillan, D., Ovalle, D., and Schaffer, A. (2025, January 13). *Police Ignore Standards after AI Facial Recognition Matches*. The Washington Post.
<https://www.washingtonpost.com/business/interactive/2025/police-artificial-intelligence-facial-recognition/>
- Mazoomdaar, J. (2019, December 28). *Delhi Police film protests, run its images through face recognition software to screen crowd*. The Indian Express.
<https://indianexpress.com/article/india/police-film-protests-run-its-images-through-face-recognition-software-to-screen-crowd-6188246/>
- Mehrotra, K. (2020, March 12). *In Lok Sabha, Amit Shah says Delhi rioters identified using facial scan*. The Indian Express.
<https://indianexpress.com/article/india/in-lok-sabha-amit-shah-says-delhi-rioters-identified-using-facial-scan-6309932/>
- Moskaleva, N. (2024). *A Brief History of Privacy*. IDURA.
<https://idura.eu/blog/history-of-privacy/>
- Nwabueze, R.N., and White, M. (2024). Privacy law and the dead – a reappraisal. *Journal of Media Law*, 16(2), 468–502.
<https://doi.org/10.1080/17577632.2024.2438395>
- Privacy. (2023, October 19). *Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/entries/privacy/>
- Rai, N. (2022). Why ethical audit matters in artificial intelligence? *AI and Ethics*, 2(1), 209–218. <https://doi.org/10.1007/s43681-021-00100-0>
- Ramiro, A., and de Queiroz, R. (2022). Cypherpunk. *Internet Policy Review*, 11(2). <https://doi.org/10.14763/2022.2.1664>
- Rhodes, C. (2025, February 13). *The Abuse of Facial Recognition Technology in the Hong Kong Protests*. UAB The University of Alabama at Birmingham: Institute for Human Rights Blog.
<https://sites.uab.edu/humanrights/2025/02/13/the-abuse-of-facial-recognition-technology-in-the-hong-kong-protests/>
- Rosen, J. (n.d.). A Twenty-First Century Framework for Digital Privacy- White Paper Series Introduction: A Twenty-First Century Framework for Digital Privacy.
- Russakovsky, O., Sadosky, A., and Ortiz, J. (2004). *Modern Cryptography: Theory and Applications*.

<https://cs.stanford.edu/people/eroberts/courses/soco/projects/2004-05/cryptography/index.html>

- Senarath, A., and Arachchilage, N.A.G. (2019). A data minimization model for embedding privacy into software systems. *Computers and Security*, 87, 101605. <https://doi.org/10.1016/J.COSE.2019.101605>
- Solove, D.J., and Richards, N.M. (2010). *Prosser's Privacy Law: A Mixed Legacy*. <http://ssrn.com/abstract=1567693>
- The Philosophy of Data Privacy: Why We're Building Programs That Miss the Point. (2025). *Governed Chaos*. <https://www.governedchaos.com/the-philosophy-of-data-privacy-why-were-building-programs-that-miss-the-point/>
- Warren, S., and Brandeis, L. (1989). The right to privacy. In *Killing the Messenger* (pp. 1–21). Columbia University Press. <https://doi.org/10.7312/gold91730-002>
- Zalnieriute, M. (2024). Power and Protest. *The Cambridge Handbook of Facial Recognition in the Modern State*, 96–111. <https://doi.org/10.1017/9781009321211.009>
- Zhang, J., and Mora, L. (2023). Nothing but symbolic: Chinese new authoritarianism, smart government, and the challenge of multi-level governance. *Government Information Quarterly*, 40(4), 101880. <https://doi.org/10.1016/J.GIQ.2023.101880>
- Zhao, H., and Liu, T. (2025). China's social credit system and the family: Punishment and collective resistance. *Economy and Society*, 54(1), 49–69. <https://doi.org/10.1080/03085147.2024.2422187>
- Zuboff, S. (2019). Surveillance capitalism and the challenge of collective action. *New Labor Forum*, 28(1), 10–29. <https://doi.org/10.1177/1095796018819461>

Extempore on Artificial Intelligence and History Education at Schools

Ashish Ranjan

University of Delhi

DOI: [10.1201/9781042018116-21](https://doi.org/10.1201/9781042018116-21)

I am remembering an incident of nearly forty-five years ago, perhaps even earlier, when as a child I was asked to write a Hindi essay on Vigyan: Vardan ya Abhishaap-science as a boon or a bane. That question continues to trouble and engage me even today, as I speak on a very contemporary and alternative theme: artificial intelligence and history education in schools. I come from a school education background. I have taught in schools during what I consider the happiest phase of my professional life, and I am presently associated with the Department of Education. In this sense, I stand at an important intersection between what happens at the grassroots level with young learners and classroom pedagogy, and what is unfolding at the college level, where students engage with history through emerging technologies.

I would like to begin by asking some basic questions, whether one approaches them as a historian or from any other disciplinary background. How do we learn? How do people learn today? And when we teach in classrooms, how do students truly learn?

There exist well-established learning theories proposed by different schools of thought; behaviorists, cognitivists, constructivists, connectivists, and pedagogists. I will not go into their detailed explanations here. Broadly, these theories suggest that learners-children as well as adults-learn through

observation, reading, writing, imitation, repetition, and lived experience. Our everyday experiences, along with the experiences of others, play a significant role in shaping our understanding. As teachers, we are expected to create meaningful learning opportunities and learning spaces, whether inside the classroom or beyond it. This brings me to another important question: how do learners understand history? While we have some clarity on how learning takes place, historical understanding requires further reflection. I would strongly argue that history is not confined to textbooks alone; history is all around us.

The larger question, which I often raise in different forums, is whether history should be understood as a product or as a process. Many of us may agree that today history is increasingly treated as a product rather than a process. Artificial intelligence further pushes us in that direction. With a single click, we receive a complete answer. For me, however, history has always been a process-one that involves inquiry, interpretation, and reflection. In this context, another question becomes crucial: when a young person engages with history, what constitutes historical understanding? This has been explained very effectively by two American historians, Thomas Andrews and Frederick Burke, in an article published in 2007 in the journal of the American Historical Association. They argued that there are five key elements-what they call the “five Cs”- that help us determine whether someone is thinking historically. The first of these is a sense of continuity and change. If I possess an understanding of continuity and change, it indicates that I have developed a sense of historicity. The second important element of historical understanding, and one that is particularly crucial in the context of artificial intelligence, is context. Everything exists within a context. This conference, this seminar, and even our presence here today is shaped by specific contexts. There are reasons why certain individuals are present and others are not. Context gives meaning to events, actions, and interpretations, and we must always remain conscious of it.

However, when we increasingly engage with AI, context often becomes blurred or weakened. AI-generated responses may provide information, but they frequently strip that information of its historical and social grounding. This loss of context is something we need to be extremely careful about. The third element relates to causality-the cause-and-effect relationship among historical phenomena. Events do not occur in isolation; they are

shaped by multiple causes and lead to multiple consequences. Closely connected to this is the fourth element, complexity. Historical realities are layered, interconnected, and often contradictory. Whether artificial intelligence simplifies this complexity or further complicates it is an important question that we must continue to examine critically.

The fifth element is contingency-the possibility that new sources may emerge and lead to entirely new interpretations. History is never final. A new document, a new narrative, or a new perspective can reshape our understanding of the past.

The larger point I am trying to make is that history has always been constructed based on evidence, what we call sources. Every historical argument must be grounded in evidence. When we construct an argument, we must also recognize that there can be more than one interpretation. To illustrate this, let me give a simple example. If I hold a pen in my hand, for me it is a writing instrument. However, for someone else, it could serve an entirely different purpose-for instance, as a hairpin. The object remains the same, but its interpretation changes. This multiplicity of meanings is central to historical understanding.

This idea was raised very clearly in the National Curriculum Framework of 2005, which posed a fundamental question to history learners: is there one past, or are there multiple pasts? The history textbooks introduced at that time described themselves as *Our Pasts*, suggesting that there are multiple narratives and interpretations of history. This recognition of plurality is extremely important. Whenever we engage with history, controversy is inevitable. Historical sources do not speak for themselves, and therefore multiple interpretations are bound to emerge. This is why careful triangulation of sources becomes essential. Historians must examine different kinds of evidence and weigh them thoughtfully before arriving at conclusions.

When we look at the engagement with history in schools, we must acknowledge that it has never been an easy task. Those of us involved in pedagogy- the art and science of teaching, know that it is grounded in strong learning theories, while also leaving space for creativity and innovation. Teaching is therefore both a science and an art, and teachers constantly negotiate between structure and imagination.

Traditionally, the most dominant pedagogical practice in history education, whether at the school level or even in higher education, has been the lecture method. From the time of Plato and Aristotle, lectures have remained one of the most popular ways of connecting with learners. Pedagogy, at its core, is about connection-about whether I can connect with my learners or audience, and whether I am able to make sense to them.

Another widely used pedagogical technique is questioning. In India especially, we love to ask questions- why this, why not that. This tradition of questioning can be traced back to Socrates. Along with this, interactive teaching methods and dialogical approaches have always been considered essential. Dialogue -whether between two people or among many-is often regarded as one of the most effective ways of learning. Even in conflict situations, such as between nations, we often say that dialogue is the only meaningful way to resolve issues. It is important to recognise that this dialogical tradition is not exclusively Western. It is deeply rooted in the Indian knowledge system as well. We see powerful examples in dialogues such as those between Yama and Nachiketa, and between Uddalaka and Shvetaketu. These dialogues demonstrate how conversation leads individuals to higher levels of thinking and cognition.

Another crucial pedagogical method in history education is the source method. Whether at the school level or in higher education, teaching through sources remains extremely important. One should not make a historical argument or statement without supporting it with sources. Ideally, multiple sources should be used. Today, even oral sources are recognised as legitimate historical evidence. Works such as Urvashi Butalia's *The Other Side of Silence* have shown how women's oral narratives were used to reconstruct the lived experiences of the Partition of India in 1947.

In addition to this, narration, storytelling, role play, and collaborative learning have long been part of history pedagogy. Comparison is another extremely important technique. I recall an advertisement that said, “Meri saree tumhari saree se zyada safed hai”; my saree is whiter than yours. It is only through comparison that we begin to understand difference and uniqueness. In history education, comparison allows us to examine how the same issue is addressed in different contexts.

For instance, I may ask how history is taught to young learners in Uzbekistan and compare it with how we teach history in India. Through

such comparisons, we can identify similarities and differences and understand what works and what does not. This comparative perspective has become increasingly important in history education across the world. Another important pedagogical approach is what E. H. Carr described in his classic 1961 definition of history, often referred to as the pendulum method. He argued that history involves a constant dialogue between the past and the present, much like the pendulum of a grandfather clock that keeps moving back and forth. As historians, we continuously interpret the past in the light of the present, and at times, the past itself appears more contentious because the present poses difficult questions.

It is within this broader pedagogical context that artificial intelligence enters the field of history education. The very term artificial intelligence is significant. The word “artificial” indicates that it is not natural or human; rather, it is created. This distinction is important, and I will return to a critique of AI later.

In India, artificial intelligence formally entered school education in the academic year 2019-20, when it was introduced as an optional subject. Subsequently, AI began to influence history and other social science subjects as well. This shift aligned closely with the emphasis on experiential learning promoted by the National Education Policy 2020 and later by the National Curriculum Framework for School Education 2023. These policy documents encourage learning beyond textbooks, and AI appears to offer tools that support such engagement. As a result, history education has gradually moved beyond strictly textbook-based teaching at both the school and college levels. One significant development has been the use of immersive historical simulations to enhance student engagement and understanding. Technologies such as augmented reality and virtual reality enable learners to experience history in ways that were previously unimaginable.

I recently visited the Prime Ministers’ Museum and Library, and one immersive activity stood out. In this experience, visitors sit in a simulated helicopter and are shown a vision of the country's development over the next twenty years. The experience is visually stimulating and emotionally powerful. I found myself filled with a sense of national pride, feeling that my country has evolved and continues to progress. Such immersive experiences demonstrate the power of AI-driven technologies. Through

these simulations, history no longer appears distant or static. Instead, learners become immersed in historical narratives, making history feel alive and immediate. This immediacy and sense of involvement are among the most striking contributions of artificial intelligence to history education.

There are numerous generative AI tools available today. I will not list all of them, but we are all familiar with platforms such as ChatGPT and Gemini. These tools have become a serious challenge for teachers, both in colleges and in schools. Students can now instantly generate answers, sources, arguments, and even creative outputs. Tasks that once required sustained intellectual effort, extensive reading, and prolonged engagement with libraries can now be completed almost instantly.

Earlier, one had to spend days, months, and sometimes years in libraries-carefully triangulating sources, correlating evidence, and analyzing materials-before arriving at a historical argument. Today, artificial intelligence makes this process appear effortless. This instant availability of information marks a significant shift in how history is accessed and produced. AI is also being used to analyse historical documents and ancient texts, particularly in higher education.

While such applications may have limited relevance at the school level, they hold considerable potential in advanced research. I am aware, for instance, of how Sanskrit has been recognised as one of the most suitable languages for computational and AI-based research. This has been possible largely due to the growing capacities of artificial intelligence. In school education as well, AI has reshaped historical projects and assessments. History projects, often carrying significant weightage, now involve data analysis and source criticism supported by digital tools. Gamification has also emerged as an important trend. Educational games, cartoons, comics, documentaries, and interactive media are increasingly used to engage learners. Many speakers have already spoken about these innovations during this symposium.

Virtual museums represent another remarkable development. With the help of AI, I can now virtually take students to Agra, Hampi, or even the Andaman and Nicobar Islands without leaving the classroom. Such experiences bring a sense of joy, beauty, and aesthetic engagement into history learning. For me, this is one of the most rewarding aspects of using AI in history pedagogy.

AI has also transformed classroom practices such as revision and assessment. If I want to conduct a quick recapitulation or quiz, I can easily use platforms like Kahoot to assess student attentiveness and understanding within minutes. Concept maps and mind maps, which were not widely used earlier, have now become powerful tools for evaluating whether learners have developed conceptual clarity and historical understanding. Assessment itself has undergone significant change due to the integration of artificial intelligence, introducing new modes of feedback and evaluation. Based on my reflections from the field, I would like to critically examine a few aspects of artificial intelligence in history education. From the perspective of economic history, AI has been particularly useful. Large databases can now be accessed and analysed with ease. I often wonder how historians such as Irfan Habib and Shireen Moosvi engaged with vast amounts of medieval data over many years. Today, much of this data is available instantly, organized into columns and datasets, largely due to AI. This accessibility is undoubtedly a positive development.

The introduction of AI in India has also contributed to more democratic participation in education. Students from marginalised backgrounds, who may have had limited access to resources earlier, can now engage with archives and scholarly materials remotely. In this sense, AI has the potential to make historical resources more inclusive and accessible.

At the same time, this optimistic view must be balanced with concerns. Reports such as UDISE+ highlight a significant digital divide in India. A substantial proportion of the population still lacks access to functional computers and reliable internet connectivity. This remains the other side of the story. At the same time, this optimistic view must be balanced with concerns. Reports such as UDISE+ highlight a significant digital divide in India. A substantial proportion of the population still lacks access to functional computers and reliable internet connectivity. This remains the other side of the story.

It is also important to remember that the intelligence exhibited by machines is not natural. It is created by humans through algorithms, computer programs, and data. Artificial intelligence does not possess human consciousness or self-awareness. This limitation becomes particularly significant in the social sciences. When dealing with social

history, issues of gender, caste, and inequality require careful and sensitive interpretation. AI must therefore be used with great caution in these areas.

Several countries, including Sweden, have chosen not to reject artificial intelligence outright but to adopt it responsibly. Their emphasis has been on ensuring that AI use remains grounded in human values and ethical considerations. Questions of bias-whether AI-generated sources are neutral or shaped by existing prejudices- must be taken seriously.

What I wish to emphasise is the urgent need for preparation. Young teachers in schools and colleges, as well as students, must be trained to handle this powerful technology responsibly. Without such preparedness, there is a real risk that artificial intelligence may lead to biased interpretations and poor historical judgment.

From Computers to Google to Artificial Intelligence: Evolution of Digital Knowledge Infrastructures

Ministry of Science and Technology, Government of India, New
Delhi

Ministry of Science and Technology, Government of India, New Delhi

Astanakulov Olim Tashtemirovich

International Islamic Academy of Uzbekistan, Tashkent, Uzbekistan

Roushan Kumar and Ankit Kumar

University of Delhi

DOI: [10.1201/9781042018116-22](https://doi.org/10.1201/9781042018116-22)

ABSTRACT: The evolution of digital knowledge infrastructures—from early computers to search engines and, most recently, artificial intelligence—reflects profound shifts in how societies perceive, produce, and distribute knowledge. When computers were first introduced, they were met with widespread apprehension, often feared as instruments of automation that would displace human labour, centralize control, and diminish human agency. Over time, however, computers were domesticated into everyday life and institutions, becoming indispensable tools for calculation, administration, research, and communication. The advent of Google marked a decisive rupture in knowledge dissemination, reorganising information access through algorithmic search, radically compressing time and space, and transforming knowledge from a scarce institutional resource into an

instantly retrievable public good. Yet this search-centric model also introduced new hierarchies based on visibility, ranking, and platform power. Artificial intelligence now represents a significant inflexion point, simultaneously celebrated for democratising knowledge production and criticised for destabilising epistemic authority. AI systems challenge traditional vertical models of expertise by enabling horizontal, participatory, and generative forms of knowledge creation, while also raising concerns about misinformation, opacity, bias, and the erosion of intellectual accountability. This article argues that rather than merely “wreaking havoc,” artificial intelligence is reconfiguring the architecture of knowledge itself, forcing a reconsideration of trust, authorship, and governance in digital epistemologies.

Keywords: Automation, Labour, Algorithmic Governance, Human–Machine Interaction, and Technological Modernity

1 METHODOLOGY

This chapter adopts a qualitative, interdisciplinary methodological approach that combines historical analysis, conceptual inquiry, and critical theory to examine artificial intelligence as a transformative yet contested phase in the evolution of digital knowledge infrastructures. Primary sources include canonical texts on political economy, philosophy of technology, and artificial intelligence, while secondary scholarship from science and technology studies (STS), media studies, and epistemology is used to contextualize contemporary debates. The chapter employs diachronic analysis to trace continuities and ruptures from mechanisation and early computing to search engines and AI, alongside comparative thematic analysis to evaluate shifting notions of labour, expertise, and epistemic authority. Interpretive textual analysis is used to critically assess key arguments, allowing the chapter to synthesize historical perspectives with contemporary concerns surrounding bias, opacity, and accountability in AI-driven knowledge systems.

1 CONCEPTUAL MODEL: SHIFTS IN LABOUR, AUTHORITY, AND KNOWLEDGE ACROSS DIGITAL STAGES

STAGE 1: Early Computers (1940s–1960s)

Labour:

- Human-dominated
- Machines assist calculation
- Clerical and computational labour augmented

Authority:

- Centralized
- State, military, scientific elites
- Expertise held by trained specialists

Knowledge:

- Scarce
- Institutional
- Codified and controlled



STAGE 2: Institutional & Personal Computing (1970s–1990s) Labour:

- Semi-automation
- Office work digitized
- Productivity enhancement

Authority:

- Professionalized
- Universities, corporations, bureaucracy

- Credential-based expertise

Knowledge:

- Digitized
- Stored and processed
- Still gatekept



STAGE 3: Internet (1990s)

Labour:

- Networked human labour
- Communication and coordination enabled
- Emergence of digital work

Authority:

- Decentralizing
- Institutions challenged by networks
- Rise of peer-to-peer exchange

Knowledge:

- Abundant
- Open access
- Distributed but unstructured



STAGE 4: Search Engines (Google Era) (2000s)

Labour:

- Cognitive labour assisted

- Information retrieval optimized
- Attention becomes labour

Authority:

- Algorithmic mediation
- Platform-based visibility
- Authority shifts from institutions to rankings

Knowledge:

- Indexed
- Ranked
- Time–space compressed



STAGE 5: Artificial Intelligence (2010s–Present)

Labour:

- Cognitive automation
- Human–machine collaboration
- Creative and analytical tasks shared

Authority:

- Hybrid and contested
- Algorithms as epistemic actors
- Institutional authority destabilized

Knowledge:

- Generative
- Probabilistic
- Dynamically produced

This model demonstrates a gradual but profound transformation from machine-assisted labour to machine-participatory cognition, from institutional authority to algorithmically mediated legitimacy, and from scarce, codified knowledge to generative and probabilistic knowledge systems. Each stage does not erase the previous one but layers new forms of power, mediation, and epistemic tension onto existing structures. Artificial intelligence thus represents not a complete rupture but an intensification of historical trends—where labour becomes increasingly cognitive, authority increasingly algorithmic, and knowledge increasingly dynamic and contested.

2 FROM EARLY COMPUTERS TO SEARCH ENGINES AND ARTIFICIAL INTELLIGENCE

The transformation of digital knowledge infrastructures marks one of the most profound intellectual and technological shifts of the past century. These infrastructures did not emerge overnight: they evolved through successive layers of innovation in computation, networking, information retrieval, and cognitive automation, each reshaping how societies create, perceive, and disseminate knowledge.

The history begins with the invention of the modern computer. Early computing machines were designed to automate calculation and data processing, heralding a departure from purely analogue methods of managing information. Historians of computing situate this transformation within a broader narrative of technological innovation and institutional investment, with the development of programmable electronic computers in the 1940s and 1950s laying the groundwork for the information age ([Campbell-Kelly, Aspray, Yost, and Tinn, 2023](#)).

The debate over technology, mechanisation, and their implications for labour and knowledge was articulated with remarkable prescience in modern India well before the advent of digital infrastructures. Mohandas Karamchand Gandhi consistently warned against large-scale mechanisation and automation, arguing that indiscriminate reliance on machines would

displace human labour, erode artisanal skills, and exacerbate social inequality—particularly in labour-abundant societies such as India.

Global reactions to technological shifts away from human labour have historically oscillated between enthusiasm, anxiety, resistance, and adaptation. These responses are neither uniform nor linear; rather, they are shaped by economic structures, cultural values, political institutions, and prevailing ideas about progress and human dignity. From the Industrial Revolution to the age of artificial intelligence, societies have repeatedly grappled with the social consequences of mechanisation and automation.

One of the earliest and most well-documented reactions emerged during the Industrial Revolution in Britain. The introduction of mechanized looms and factory-based production in the late eighteenth and early nineteenth centuries provoked fierce resistance from skilled artisans whose livelihoods were threatened. The Luddite movement (1811–1816), often mischaracterized as irrational technophobia, was in fact a targeted protest against the social organization of production rather than machinery per se. As historians have shown, Luddites opposed the use of machines that undermined customary labour relations and eroded worker autonomy ([Hobsbawm, 1962](#); [Thompson, 1963](#)). Their actions reflected a broader fear that mechanization would devalue human skill and concentrate economic power in the hands of industrial capitalists.

In continental Europe, reactions were similarly ambivalent. Karl Marx famously analyzed machinery as a double-edged force: while technological advancement increased productive capacity, it also alienated workers from their labour and rendered them increasingly replaceable. In *Capital*, Marx argued that machines, under capitalist relations, functioned as instruments of domination, transforming workers into appendages of the industrial system ([Marx, 1867/1990](#)). This critique influenced labour movements across Europe, shaping demands for labour rights, shorter working hours, and social welfare mechanisms to mitigate technological displacement.

In the United States, the response to mechanization was more closely tied to narratives of progress and frontier expansion. While industrial automation displaced certain forms of labour, it also generated new occupations and industries, fostering a belief in technological optimism. Joseph Schumpeter later conceptualized this process as “creative destruction,” wherein innovation continuously dismantles old economic

structures while creating new ones ([Schumpeter, 1942](#)). Nevertheless, even in the American context, periods of rapid technological change—such as the automation of manufacturing in the mid-twentieth century—provoked widespread fears of technological unemployment ([Noble, 1984](#)).

In colonial and postcolonial societies, reactions to mechanization were deeply entangled with experiences of imperialism. In India, Mahatma Gandhi articulated a powerful moral critique of industrial machinery, warning that large-scale mechanization would displace labour in a labour-surplus society and destroy village-based economies. Gandhi viewed machines not as neutral tools but as social forces capable of reshaping human relations in ethically destructive ways ([Gandhi, 1909/2010](#)). His resistance reflected broader anti-colonial concerns that Western industrial technologies served imperial interests rather than indigenous well-being.

Following independence, however, many postcolonial states adopted a contrasting stance. Leaders such as Jawaharlal Nehru embraced technology and industrialization as essential to national development, seeing large-scale infrastructure and technical institutions as engines of modernity and self-reliance ([Nehru, 1946/1999](#)). This divergence illustrates how reactions to technology are conditioned by political objectives: while Gandhi emphasized moral economy and employment, Nehru prioritized scientific capacity and state-led modernization.

In the late twentieth century, the rise of computerization and digital automation generated a new wave of global concern. Scholars noted that unlike earlier mechanization, digital technologies threatened not only manual labour but also clerical, managerial, and professional work. Harry Braverman's influential analysis demonstrated how automation and management technologies systematically deskilled labour, reducing worker control even in white-collar occupations ([Braverman, 1974](#)). These concerns intensified with the spread of information technologies across global economies.

In the contemporary era, artificial intelligence has amplified historical anxieties about technological displacement. AI systems now perform tasks involving analysis, prediction, and language—functions traditionally associated with human cognition. Economists and sociologists have warned that AI-driven automation could exacerbate inequality if institutional safeguards fail to accompany technological adoption ([Autor, 2015](#)). Public

reactions today echo earlier historical patterns: optimism about efficiency and innovation coexists with fears of job loss, surveillance, and the erosion of human agency.

Across historical contexts, a recurring pattern emerges: societies rarely reject technology outright, but they resist technological change when it threatens livelihoods, social equity, or moral values. Acceptance tends to follow when institutions adapt through education, labour protections, and redistributive mechanisms. Thus, global reactions to the technological displacement of human labour reveal that technology is not merely a technical phenomenon—it is a social process negotiated through history, power, and collective choice.

For Gandhi, machines were not inherently evil, but their uncritical adoption in the absence of ethical and social considerations posed a fundamental threat to human dignity and self-reliance ([Gandhi, 1938/2001](#)). His critique of industrial civilisation was rooted in the conviction that technology must serve human ends rather than reorganize society around mechanical efficiency. Gandhi feared that mechanisation, when divorced from moral restraint, would generate unemployment on a mass scale and undermine village-based economies that sustained India's social fabric ([Gandhi, 1909/2010](#)).

Gandhi's skepticism toward machinery extended beyond economics into epistemology. He questioned the modern tendency to equate progress with technological advancement, arguing instead for a knowledge system grounded in ethical judgment, manual skill, and lived experience. In this sense, Gandhi's vision anticipated contemporary critiques of digital knowledge infrastructures that caution against over-automation and the displacement of human agency by algorithmic systems. His insistence that technology remain subordinate to social purpose resonates with present concerns about artificial intelligence replacing human judgment, labour, and creativity rather than augmenting them.

In contrast, the post-independence Indian state under Jawaharlal Nehru embraced technology as an indispensable instrument of national development and modernization. Nehru viewed science, technology, and large-scale industrialization as essential to overcoming poverty, underdevelopment, and colonial stagnation. Famously describing dams, laboratories, and scientific institutions as the “temples of modern India,”

Nehru articulated a vision in which technological infrastructure would function as the backbone of a rational, secular, and progressive nation-state ([Nehru, 1946/1999](#)). Unlike Gandhi, Nehru perceived machines not as threats to human labour but as tools capable of expanding productivity, knowledge, and national capacity when guided by democratic planning.

Central to Nehru's technological vision was the promotion of technical education and scientific institutions. He championed the establishment of institutes of higher learning—particularly in science, engineering, and technology—as engines of knowledge production and innovation. For Nehru, education in science and technology was not merely vocational but civilizational, enabling citizens to cultivate what he called the “scientific temper”—a mode of thinking grounded in rational inquiry, evidence, and openness to change ([Nehru, 1989](#)). This emphasis laid the intellectual foundations for India's later engagement with digital technologies, information systems, and, eventually, artificial intelligence.

Read together, Gandhi and Nehru represent two contrasting yet complementary strands in India's engagement with technology and knowledge infrastructures. Gandhi foregrounded ethical restraint, human labour, and social responsibility, while Nehru emphasized institutional capacity, scientific rationality, and technological modernization. Contemporary debates on digital knowledge infrastructures—ranging from search engines to artificial intelligence—echo this unresolved tension. As AI systems increasingly mediate knowledge production and distribution, the Gandhian concern over displacement and moral accountability intersects with the Nehruvian faith in institutionalized science and technical expertise. The evolution of digital knowledge infrastructures thus reflects not merely technological progress, but a deeper philosophical negotiation between human values and machine-mediated knowledge systems.

These parallel currents of caution and curiosity did not evolve in isolation. While social critics, political leaders, and labour movements grappled with the disruptive consequences of mechanisation and automation, scientific inquiry continued to advance along its own trajectory. The tension between ethical restraint and technological ambition became a defining feature of modernity itself. As societies negotiated the social costs of machines, researchers simultaneously pursued the possibility of extending human cognition through computational means. It was within this

historical interplay—between social apprehension and scientific aspiration—that artificial intelligence emerged as a serious intellectual project, linking earlier debates on labour displacement with new questions about knowledge, reasoning, and the role of machines in human decision-making ([Nilsson, 2024](#); [Crevier, 1993](#)).

With the proliferation of personal computers and the commercialization of networking technologies in the late 20th century, digital infrastructure expanded into global communication systems. The emergence of the World Wide Web transformed the internet from a specialized research tool into a universal platform for information exchange. Search engines became indispensable to navigating this vast digital corpus, evolving from simple index-based tools to complex algorithmic systems that rank and retrieve information at scale ([Kathuria, Nagpal, and Duhan, 2016](#)).

Search technologies themselves represent a pivotal shift in knowledge infrastructures: they externalized aspects of human memory and judgment, enabling users to query and access distributed knowledge repositories instantly. This transition restructured both epistemic authority and public access to information, democratizing retrieval while simultaneously introducing new challenges around relevance, bias, and digital literacy. Over time, search engines integrated increasingly sophisticated techniques—including link analysis, user profiling, and semantic understanding—to enhance precision and contextual relevance.

The most recent inflection point in this evolution involves artificial intelligence. AI technologies—particularly machine learning and natural language processing—extend the scope of digital knowledge systems beyond retrieval toward generation. AI models can synthesize information, infer patterns, and respond to complex queries in ways that approximate human reasoning. This shift blurs the boundary between passive information retrieval and active knowledge production, with implications for scholarship, decision-making, and public discourse ([Russell and Norvig, 2020](#)).

Artificial intelligence not only operationalizes knowledge but also transforms the very infrastructures through which knowledge is organized. AI systems underpin modern search enhancements, recommendation engines, and automated indexing, reshaping how users encounter and interpret digital content. Critically, the integration of AI into knowledge

infrastructures raises questions about epistemic authority, algorithmic transparency, and the social distribution of information power — issues that are increasingly central to debates in digital culture and policy.

Taken together, these developments illustrate how digital knowledge infrastructures have evolved from mechanized computation to global networks and intelligent systems. Each phase reflects a broader shift in epistemological frameworks: from localized expertise to distributed information ecosystems, and now toward hybrid human–machine intelligences that co-produce knowledge. This evolution not only reflects technological progression but also deeper social transformations in how knowledge is produced, validated, and shared across collective actors.

3 THE TRANSFORMATION OF DIGITAL KNOWLEDGE INFRASTRUCTURES

Over time, computers underwent a profound process of domestication, gradually integrating themselves into everyday life, organizational routines, and institutional structures. What began as large, centralized machines designed for specialized scientific and military calculations evolved into indispensable tools embedded within the quotidian functioning of modern societies. Early computers were primarily deployed within state agencies, universities, and large corporations, where they were used for numerical computation, census operations, payroll management, and scientific research. This transition marked a shift not only in technological capability but also in social imagination, as computers were slowly normalized as routine instruments of governance and knowledge work ([Campbell-Kelly *et al.*, 2023](#), 83–112).

By the late twentieth century, computers had become integral to administration, education, research, and communication. Bureaucratic systems increasingly relied on digital databases; universities reorganized knowledge production around computational tools; and communication migrated from paper-based correspondence to electronic formats. Importantly, computerization did not merely automate existing practices but restructured them at a deeper epistemic level. Shoshana Zuboff in his study

demonstrates that information technologies transformed how organizations *knew* their environments, introducing new regimes of visibility, monitoring, and decision-making ([Zuboff, 1988](#), 9–15, 351–356). Knowledge, once embedded primarily in human judgment and institutional memory, increasingly became codified, stored, and processed in machine-readable forms.

The emergence of the Internet constituted the next major transformation in digital knowledge infrastructures. While computers initially concentrated informational power within specific institutions, the Internet dissolved spatial, organizational, and temporal boundaries by networking machines globally. Manuel Castells conceptualized this shift as the rise of the “network society,” wherein flows of information supplanted physical location as the primary organizing principle of social and economic life ([Castells, 1996](#), 469–472). The Internet radically altered communication practices, enabling near-instantaneous exchange across vast distances and fostering new modes of collaboration, authorship, and dissemination.

Within this expanding digital ecosystem, search engines—and most notably Google—introduced a decisive rupture in the organization of knowledge access. Whereas early digital networks expanded availability, search engines restructured intelligibility itself. Google's central innovation lay in algorithmic ranking: information was no longer encountered primarily through curated institutional channels but through automated relevance hierarchies determined by links, popularity, and user behaviour. Google reshaped epistemic authority by privileging speed, scale, and algorithmic relevance over traditional criteria of expertise and editorial judgment ([Vaidhyanathan, 2011](#), 56–61).

This search-centric model radically compressed time and space. Queries that once demanded prolonged engagement with archives, libraries, or experts could now be resolved within seconds. In this process, knowledge was transformed from a scarce, institutionally mediated resource into an apparently abundant and instantly retrievable public good. However, this abundance also altered cognitive habits as reliance on search engines encourages fragmented attention and shallow engagement, privileging retrieval over reflection and accumulation over comprehension ([Carr, 2010](#), 115–121).

Yet this democratization of access simultaneously produced new hierarchies. Search engines did not eliminate power; they redistributed it through algorithmic visibility. What appeared on the first page of results acquired epistemic legitimacy, while what remained buried effectively vanished from public consciousness. Platforms such as Google as “custodians of public knowledge,” exercising immense, often opaque authority over what is visible, credible, and consequential in the digital public sphere ([Gillespie, 2018](#), 67–75).

Moreover, search engines reshaped not only access to knowledge but also its production. Content creators increasingly optimized information to align with algorithmic logics, subtly subordinating epistemic values to metrics of visibility and ranking. This transformation embedded knowledge production within commercial infrastructures governed by proprietary algorithms, introducing tensions between scholarly rigour and platform-driven incentives ([Vaidhyanathan, 2011](#), 182–188).

Viewed historically, the evolution from domesticated computers to networked search engines represents a fundamental reorganization of digital knowledge infrastructures. Computers mechanized calculation and administration; the Internet globalized communication; and search engines algorithmized access to knowledge itself. Each stage deepened the entanglement of technology with human cognition and institutional life. Together, these developments prepared the ground for the contemporary rise of artificial intelligence, where machines increasingly move beyond retrieval toward interpretation, synthesis, and generation of knowledge. The trajectory underscores a central insight: technological progress in knowledge systems is never neutral—it reshapes how knowledge is structured, valued, and governed.

4 ARTIFICIAL INTELLIGENCE: ADVENT OF A NEW ERA OR THE RETURN OF AN OLD BOGEY?

Artificial intelligence occupies an ambivalent position in contemporary debates on technology and society. It is simultaneously celebrated as a transformative force inaugurating a new phase of human–machine

collaboration and viewed with apprehension as a renewed manifestation of long-standing anxieties concerning automation, labour displacement, and the erosion of human agency. This duality invites a historically grounded interrogation of whether AI represents a radical epistemic rupture or the re-emergence of familiar technological fears in an advanced computational form.

Concerns about intelligent machines are deeply embedded in the history of modernity. From the mechanisation of production during the Industrial Revolution to the twentieth-century rise of automation and cybernetics, technological innovation has repeatedly provoked fears of unemployment, alienation, and the concentration of power. Marx's analysis of machinery under industrial capitalism demonstrated how technological progress often intensified exploitation by subordinating human labour to the imperatives of capital rather than liberating it (Marx, 1867/1976, 492–508). Similarly, Norbert Wiener warned that automated and feedback-driven systems, if left without ethical and political oversight, could generate widespread social dislocation and undermine human responsibility (Wiener, 1950, 162–181).

Artificial intelligence inherits these historical anxieties while simultaneously intensifying them. Unlike earlier machines designed primarily to augment physical labour or accelerate calculation, AI systems increasingly perform tasks associated with cognition, including pattern recognition, decision-making, language generation, and creative production. This qualitative shift positions AI not merely as an instrumental technology but as an epistemic actor participating in processes of knowledge production. Early pioneers of artificial intelligence explicitly envisaged machines capable of simulating aspects of human reasoning, grounding this ambition in formal logic and symbolic computation (McCarthy et al., 1955, 1–2; Newell and Simon, 1976, 113–116). Over time, these theoretical foundations evolved into operational systems that now shape research, governance, education, and everyday communication (Nilsson, 2010, 12–18).

Within the longer trajectory of digital knowledge infrastructures, AI constitutes a decisive inflection point. Whereas computers institutionalised calculation and administration, and the internet—followed by search engines such as Google—revolutionised access to information, AI transforms the processes through which knowledge itself is generated,

synthesised, and interpreted. Rather than merely facilitating retrieval, AI systems actively shape meaning and inference. Scholars argue that this development challenges traditional vertical models of expertise by enabling more horizontal, participatory, and generative forms of knowledge creation, thereby redistributing epistemic authority beyond established institutions ([Floridi, 2014](#), 85–92; [Russell and Norvig, 2020](#), 1–5).

At the same time, the apparent democratisation of knowledge production enabled by AI introduces significant epistemic risks. Machine-learning systems and natural language models operate through complex computational processes that are frequently opaque, even to their designers. This opacity complicates established norms of accountability, authorship, and justification that underpin scholarly and institutional knowledge practices. Algorithmic opacity undermines transparency and challenges the conditions under which claims to knowledge can be explained, evaluated, and contested ([Burrell, 2016](#), 3–7).

Moreover, AI systems often reproduce and amplify biases embedded in their training data and institutional contexts. Rather than functioning as neutral arbiters, they reflect and reinforce existing social hierarchies and power asymmetries, shaping whose knowledge is rendered visible, legitimate, or marginal. Analysis of algorithmic systems illustrates how automated knowledge infrastructures can perpetuate racial, gendered, and cultural biases while maintaining an appearance of objectivity ([Noble, 2018](#), 10–15). In this respect, AI mirrors earlier technological regimes in which innovation promised efficiency and progress while consolidating power within dominant actors (Ellul, 1964, 133–147).

The generative capacities of AI further intensify concerns regarding misinformation and epistemic destabilisation. Unlike search engines that primarily organise and rank existing content, AI systems can generate novel yet plausible outputs that blur distinctions between knowledge, interpretation, and fabrication. This erosion of epistemic boundaries raises profound questions about truth, responsibility, and intellectual authority. Philosophical reflections on thoughtlessness and moral responsibility in technologically mediated societies acquire renewed relevance in an era where judgment and decision-making are increasingly delegated to algorithmic systems (Arendt, 1958, 3–6).

Taken together, artificial intelligence both extends and disrupts the historical evolution of digital knowledge infrastructures. It expands access, participation, and creative capacity while simultaneously unsettling established epistemic hierarchies and accountability structures. As with earlier technological transformations, the significance of AI lies not solely in its technical capabilities but in the social, political, and ethical frameworks through which it is governed and embedded. Whether AI ultimately signals the advent of a genuinely new era or the return of an old technological bogey will depend on how societies negotiate questions of power, responsibility, and trust in an increasingly algorithmic world.

5 CONCLUSION

The evolution of digital knowledge infrastructures, from the earliest mechanical calculators to the sophisticated artificial intelligence systems of today, reflects a profound reconfiguration of how societies generate, mediate, and contest knowledge. Historical analysis demonstrates that every major technological advance—from the Industrial Revolution through the advent of electronic computers and the rise of the Internet—has been accompanied by both transformative opportunities and persistent anxieties about labour, authority, and social equity (Marx, 1867/1976; Wiener, 1950). AI, as the latest stage in this trajectory, represents both the culmination and intensification of these long-standing dynamics. Unlike previous technologies, which primarily augmented human physical or cognitive effort, contemporary AI systems participate actively in knowledge production, performing tasks that require decision-making, pattern recognition, language generation, and even creative synthesis ([Russell and Norvig, 2020](#), 1–5; Nilsson, 2010, 12–18). This marks a significant epistemic shift: knowledge is no longer the exclusive preserve of credentialed institutions or human specialists but emerges through human–machine collaboration, algorithmic mediation, and distributed processes of validation ([Floridi, 2014](#), 85–92; [Jasanoff, 2004](#), 14–18).

At the same time, AI revives enduring social and ethical concerns. Historically, fears regarding automation and mechanization centred on the

displacement of labour, the concentration of power, and the alienation of human activity (Ellul, 1964, 133–147). AI intensifies these anxieties, particularly as cognitive tasks—once considered uniquely human—are increasingly outsourced to machines. This raises fundamental questions about agency, accountability, and intellectual responsibility (Arendt, 1958, 3–6). Algorithmic opacity, the reproduction of bias in training data, and the concentration of technological power in a few global corporations exemplify the persistence of the “old bogey” even as the capacities of AI create unprecedented opportunities for democratization and knowledge generation (O’Neil, 2016, 3–7; [Noble, 2018](#), 10–15; Pasquale, 2015). In this duality, AI embodies both the promise of a new epistemic era and the risk of perpetuating historical inequalities.

The trajectory of digital knowledge infrastructures also demonstrates the critical role of social, political, and institutional frameworks in shaping technological outcomes. While computers, the Internet, and search engines transformed access to information, the emergence of AI underscores the importance of governance, ethical design, and participatory oversight in determining whether technological power enhances or undermines collective well-being ([Floridi, 2014](#), 87–92). The democratization of knowledge facilitated by AI is contingent upon careful regulation, transparency, and the cultivation of critical digital literacy. Absent these mechanisms, AI risks consolidating epistemic authority within algorithmic black boxes, creating new hierarchies based on visibility, ranking, and platform dominance, rather than on the quality, validity, or ethical grounding of knowledge itself ([Burrell, 2016](#), 3–7; [Russell and Norvig, 2020](#), 1–5).

From a labour perspective, the evolution of digital infrastructures highlights the shifting boundary between human and machine activity. In early computing, machines primarily augmented human physical and cognitive labour, with authority residing in institutional hierarchies (Marx, 1867/1976). The Internet and search engines facilitated networked labour, distributed knowledge, and new forms of peer-to-peer authority. AI, by contrast, actively participates in cognitive processes, reshaping what constitutes human expertise and knowledge work. Human labour is now increasingly collaborative, creative, and supervisory, while machines assume tasks traditionally associated with judgment and interpretation. This

hybrid model presents both possibilities for enhanced productivity and innovation, and challenges for ethical oversight and the maintenance of epistemic responsibility ([Russell and Norvig, 2020](#), pp. 1–5; Nilsson, 2010, 12–18).

Epistemologically, AI transforms knowledge from a relatively static, codified resource into a dynamic, generative, and probabilistic phenomenon. Knowledge is no longer purely stored or retrieved; it is continuously synthesized, modeled, and, in some cases, co-created with AI systems. This generates novel forms of authority and accountability, demanding new frameworks for assessing credibility, verifying claims, and maintaining scholarly and societal trust ([Floridi, 2014](#), 87–92; [Jasanoff, 2004](#), 14–18). At the same time, it intensifies anxieties surrounding misinformation, algorithmic bias, and the erosion of traditional mechanisms for establishing truth and expertise. These developments underscore that AI is not merely a technological phenomenon but a socio-epistemic one: its impact on society depends on how humans interpret, govern, and integrate it into existing knowledge ecosystems.

In sum, artificial intelligence represents a complex inflection point in the history of digital knowledge infrastructures. It embodies the dual potential to democratize knowledge production while simultaneously destabilizing long-established epistemic hierarchies. Whether AI is regarded as the advent of a genuinely new era or as the return of an old technological bogey depends on the social, political, and ethical frameworks that guide its development, deployment, and integration into human systems. Its significance lies not solely in technical capacity but in the governance of knowledge, labour, and authority in an increasingly algorithmically mediated world. As history demonstrates, technological power alone does not determine societal outcomes; human choices, oversight, and critical engagement remain decisive in shaping whether AI becomes a tool of emancipation or a new instrument of inequity (Ellul, 1964, 133–147; Marx, 1867/1976, 492–508; Wiener, 1950, 162–181).

Ultimately, the study of AI within the broader arc of digital knowledge infrastructures invites a reflective and historically informed approach. It challenges scholars, policymakers, and society at large to navigate the delicate balance between technological innovation and ethical responsibility, between expanding cognitive capacities and preserving

human agency. This chapter demonstrates that AI's promise and peril are inextricably intertwined: only through careful governance, critical literacy, and active engagement with both technological and social dimensions can its potential be harnessed in ways that foster equitable knowledge creation, sustainable labour practices, and robust epistemic authority.

REFERENCES

- Autor, D.H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30. [↗](#)
- Braverman, H. (1974). *Labor and monopoly capital: The degradation of work in the twentieth century*. Monthly Review Press. [↗](#)
- Burrell, J. (2016). How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. [↗](#)
- Campbell-Kelly, M., Aspray, W., Yost, J.R., and Tinn, H. (2023). *Computer: A History of the Information Machine* (3rd edn ed.). Routledge. [↗](#)
- Carr, N. (2010). *The Shallows: What the Internet is Doing to our Brains*. W. W. Norton & Company. [↗](#)
- Castells, M. (1996). *The Rise of the Network Society*. Blackwell. [↗](#)
- Crevier, D. (1993). *AI: The Tumultuous History of the Search for Artificial Intelligence*. BasicBooks. [↗](#)
- Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press. [↗](#)
- Gandhi, M.K. (2001). *The Collected Works of Mahatma Gandhi* (Vols. 1–100). Publications Division, Government of India. (Original writings, various years) [↗](#)
- Gandhi, M.K. (2010). *Hind Swaraj and other writings* (A.J. Parel, Ed.). Cambridge University Press. (Original work published 1909) [↗](#)
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press. [↗](#)
- Habib, I. (2018). *Indian nationalism: The Gandhian legacy*. Tulika Books.

- Harari, Y.N. (2024). *Nexus: A Brief History of Information Networks from the Stone Age to AI*. Fern Press/Penguin Random House.
- Hobsbawm, E.J. (1962). *The Age of Revolution: 1789–1848*. Vintage. [!\[\]\(83eb2aa26b610eb6a9dca7cf4702d681_img.jpg\)](#)
- Jasanoff, S. (2004). *States of Knowledge: The Co-production of Science and Social Order*. Routledge. [!\[\]\(94dfacbf937cdd7da4837a6fcd8fc785_img.jpg\)](#)
- Kathuria, M., Nagpal, C.K., & Duhan, N. (2016). Journey of web search engines: Milestones, challenges & innovations. *International Journal of Information Technology and Computer Science*, 12, 47–58. [!\[\]\(dae8c3c5fa7c80febd6526a5e8a853bf_img.jpg\)](#)
- Khanna, R. (2016). *Dark Continents: Psychoanalysis and Colonialism*. Duke University Press. (For interpretive context on technology, modernity, and postcolonial thought)
- Marx, K. (1990). *Capital: Volume I* (B. Fowkes, Trans.). Penguin Classics. (Original work published 1867) [!\[\]\(8f38ab9775d1331a4e1fd6648d0a83f1_img.jpg\)](#)
- Nehru, J. (1989). *Jawaharlal Nehru: Selected works* (2nd edn Series, Vols. 1–15). Oxford University Press. [!\[\]\(5e48b3241d711ef916255d822ab3415f_img.jpg\)](#)
- Nehru, J. (1999). *The discovery of India*. Oxford University Press. (Original work published 1946) [!\[\]\(c4c3604751fde0df44855086d7798e30_img.jpg\)](#)
- Nilsson, N.J. (2024). *The Quest for Artificial Intelligence: A history of ideas and achievements*. Cambridge University Press. [!\[\]\(a16ee0329f478adb49fd7876490e96fe_img.jpg\)](#)
- Noble, D.F. (1984). *Forces of Production: A Social History of Industrial Automation*. Oxford University Press. [!\[\]\(a8a494e883c31c6f4a74584cc935a790_img.jpg\)](#)
- Noble, S.U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press. [!\[\]\(8a50a875cd94cd57f6a7348a4d34b45f_img.jpg\)](#)
- Russell, S., and Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th edn ed.). Prentice Hall. [!\[\]\(7ddc0498d10972f0f157498fccf65370_img.jpg\)](#)
- Schumpeter, J.A. (1942). *Capitalism, Socialism and Democracy*. Harper & Brothers. [!\[\]\(1481f3a2b8d8d64dbbf1eb2242b57620_img.jpg\)](#)
- Thompson, E.P. (1963). *The Making of the English Working Class*. Vintage. [!\[\]\(1d12b5523dd7503f9f6b056602d4b0bb_img.jpg\)](#)
- Vaidhyanathan, S. (2011). *The Googlization of Everything (and why we should worry)*. University of California Press. [!\[\]\(a3cf9896eece75d50b7dee194989c9bc_img.jpg\)](#)
- Zuboff, S. (1988). *In the Age of the Smart Machine: The Future of Work and Power*. Basic Books. [!\[\]\(26147612c3694ee09d020cac0faae7d9_img.jpg\)](#)

From Gurukul to Digital Learning and the Evolution of Education with Modern technology and Artificial Intelligence

Vikram Raj

Jawaharlal Nehru University

Ishwar Dan

University of Delhi

DOI: [10.1201/9781042018116-23](https://doi.org/10.1201/9781042018116-23)

ABSTRACT: Education has undergone through enormous changes with evolution of time and innovation is adding a new dimension in the field of education. With the concept of digital learning and AI, Edtech revolution has brought the world closer in academics. With gradual change in technology, advancement could be noticed in diverse structures of education. Therefore, it has no longer limited to the traditional type of classroom. Some of the shining examples like use of Smart board, AI-driven visual representation and local language made easier to understand the complex problems. Digital learning has revolutionized the way of learning, teaching and communication in education. This deep dive into the new technologies that are driving the Edtech revolution needs to put emphasis on their impact on the future of learning also. From traditional classrooms to AI-powered learning tools, educators and institutions are using this modern technology to improve educational outcomes, personalize learning experiences and expand access to quality type of education. India

AI mission shows a bold vision to make the nation is emerging as a frontrunner in the field of AI-powered education.

Keywords: Edtech, Digital learning, AI, Education, Classroom, Gurukul etc

1 METHODOLOGY

This study adopts a qualitative, historical–descriptive methodology, drawing on secondary literature and policy documents to trace the evolution of education from the Gurukul system to digital and AI-enabled learning. A comparative analytical approach is used to assess how modern technologies reshape pedagogical practices and educational access in India.

2 INTRODUCTION

Over the time Indian education system has been transformed and brought major changes in teaching and learning methods. Earlier Gurukul mode of learning was the prominent feature but now modern technology-based learning came to foreground. This highlights the adaptability of India's changing social needs, such as growth of technological innovations and the consequent digital or virtual mode of learning. In this chapter we would be exploring the transformation of teaching methods in India, and how it evolves from the Gurukul system to the digital environment. Coming to Gurukul system its roots can be traced to ancient times “where the students (shishyas) lived along with their teachers (Gurus) and education was imparted at the Guru's home, which was known as Gurukul” ([Shanwal, 2023](#)). In Gurukul, education combined both philosophical percepts along with the intellectual and moral growth. Mode of learning was based on oral traditions; focus was on Vedas, Upanishads and Sanskritic education. If Gurukul system has been reintroduced in the modern India, we can create a generation based on good value system and it is possible to arrive at a holistic approach by combining traditional learning with the scientific

digital technology. Historian Radha Kumud Mookerji described Gurukul was less of formal and more a holistic lifestyle, where education was deeply rooted in Guru's daily practice and the student's interaction with their nature. In the following chapter we would explore the Gurukul schooling methods and how it transformed with the advent of modern education, digital technology. Further how AI would enhance the present education and what are the policy recommendations for supporting its effective implementation would be explored. AI-enabled learning and the preparation of AI workforces are also new parts of different nation's educational policies. In India, for example, NEP 2020 framework states that “New technologies involving artificial intelligence, machine learning, block chains, smart boards, handheld computing devices, adaptive computer testing for student development, and other forms of educational software and hardware will not just change what students learn in the classroom but how they learn” ([Government of India, 2020](#),54). Use of digital technology in India would revolutionize the modern education system and this would bring flexibility and convenient in teaching and learning methods.

3 INDIAN EDUCATION SYSTEM IN TRANSITION

In Gurukul system, guru and shisya were considered equal and used to live together and their relationship was so sacred that no fees charge. In return students have to give gurudakshina which was the token of gratitude ([Pareek, 2021](#)). It was in this way the students acknowledged, thanked and respect their gurus whom they considered their spiritual guide. The best example of relationship between guru-students can be noted from Dronacharya and Eklavya. “Gurukula is a Sanskrit word which means guru, (teacher or master) and Kula (family or home) ([Shanwal, 2023](#)). In Hindu tradition, relationship between guru-shishya is considered to be holy and sacred. “Texts like vedas, Upanishads, the epics are important and upheld the true values of Indian society” (Shelly, 2015). These texts impart knowledge to the students. India is renowned for its rich cultural and traditional heritage, recognized globally. It is the birthplace of great scholars like Aryabhata, Panini, and Chanakya, and home to prestigious ancient

institutions such as Nalanda and Takshashila. The country has also contributed significantly to global knowledge through many works like the Arthashastra, which offered deep insights into “social and political sciences, and the invention of zero” ([Madhekar, 2020](#)). The Mahabharata contains Sanskrit verse which highlight about the Guru in the Gurukul system of education:

गरुन्धरन्धनागरुश्चक्षुरचक्षुषाम्।

गुरुः पिताच माताच सर्वशास्त्रेषु पात्रकम्॥

(Sukthankar, 1947-1966)

This highlights the comprehensive role of a guru in one's life the guru is not just a teacher, but also a parent, friend, and good guide. Those who live by dharma (righteousness), the guru imparts wisdom, help, and support, and this leads students to enlightenment and ethical way of living.

In India, university like Nalanda and few others were better known globally and has been described for their tremendous work contribution in diverse fields such as humanism to science. This university has also been mentioned by the foreign writers like Greeks and Chinese traveller. “Gurukul emphasis on self-discipline, humanism and spiritual development, has inspired the modern education system and beyond” ([Selvamani, 2019](#)). Gurus in those days had profound knowledge and devoted time to provide highest quality education to the disciples and try to bring talents hidden in them. Students were obedient and have all the good qualities. They used to learn with curiosity, dedication and practice. Vedic education used to be given orally rather than written form. Education consists of three stages; the first stage is Shravana (listening), which focuses on gaining knowledge through the listening of the shrutis. The second stage is Manana (contemplation), where learners think critically, analyse information, and draw conclusions. The third phase is Nididhyasana, where students

implement the knowledge, they have acquired in practical situations. During the Vedic era, the foundation of education was predominantly centred around the Vedas. Vedic education encompassed accurate pronunciation and recitation of the Veda. Knowledge of medicine was also part of the curriculum. According P. Dharampal, the Gurukul system's decline began during British colonial rule in the 19th century. Colonial administrators, such as Thomas Babington Macaulay, dismissed indigenous education as “backward” and imposed a Western-style, examination-centric model through the 1835 English Education Act (1983). The Gurukul system, originating in ancient India, emphasized holistic development by nurturing physical, mental, and spiritual growth. The strong bond which is created between teacher and student, supported by a personalized curriculum that incorporated moral education, practical skills, and spiritual scriptures. As the Gurukul system thrived in ancient India, the arrival of new cultures and rulers gradually transformed the country's approach to education. During the Medieval period, this evolution took on new dimensions, as Islamic traditions deeply influenced the patterns of learning and knowledge in India.

During the medieval times, Muslim religious mode of education was more prominent; it differed from earlier Buddhist and Brahmanical education. Islamic influence on education was more apparent. Apart, from the knowledge transmission, teachings of Islam were also imparted. Thus, religion formed the basis of the curriculum. Along, with making students aware of the important aspects of life, religious values, devotion were also instilled. With the Bismillah ceremony formal education was often started, it can be compared with the upanayana ritual of Ancient India. Maktabas were the institutions for primary education and Madrasas were for advanced learning. Royal, support has greatly helped for the spreading of education (Jafri and Zaheer, 2012). The Mughal emperors had played an important role in shaping education system. Pathashalas, vidhyappethas, maktabas and madrasas were served as the centers for learning ([Biswas, 2016](#)). In medieval India Urdu had gained popularity, it was developed through cultural exchange between the Persian and Hindi languages and it set as a unique example of cosmopolitan culture. The Mughal emperor Humayun introduced subjects like Mathematics, astronomy, and geography. In Delhi madrasas during Akbar's rule there was significant progress in education

which had surpassed Tughlaq dynasty's contributions and had a long-lasting influence on the society. Thus, the interaction between the earlier existing indigenous practices and the new Islamic traditions brought by Delhi Sultanates and Mughals had laid foundation for subsequent transformations in India's education system. With the advent of Britishers new institutions and modes of learning were being introduced in Indian soil, which had paved the way for pre and post independent period developments.

During the modern times, National Education Policy of 1968 focused on establishing a uniform education system across the country. The policies of 1986 and 1992, on the other hand, emphasized ensuring both quality and accessibility in India's education system. After 34 years, the NEP 2020 introduced major reforms, focusing on holistic learning and a multi-disciplinary approach ([Gopalkrishnan, 2023](#)). Education can be inferred as the transfer of culture from the present to the future generations, with the aim of enabling individuals to live effectively and enhancing productivity respectively. During the British period significant changes were brought in education system. It was initiated by Christian missionaries and gradually this resulted in development of four types of educational institutions such as: church based primary schools, colleges for higher education, theological institutions, vocational and agricultural institutions. St. Francis Xavier was a notable Christian missionary; he had followed teachings of Rober Nobilli and worked for the spread of Christian missionary and he had even invited Brahmins to accept his teachings ([Joshi, 2020](#)). During the pre-independent period India has three main systems of education such as

- a. The Hindu system, where students were taught Sanskrit before moving on to subjects such as the Vedas, philosophy, astronomy, and grammar, primarily led by Brahmins.
- b. The Muslim system, where education was imparted in maktabas and madrasas, run by the Mohammedans.
- c. The Christian missionary system, which focused on teaching Christianity.

Thomas Babington Macaulay, a British historian and politician who has greater importance in forming educational structures of India. In his Minute on Education, he strongly advocated for English to be adopted as the

medium of instruction. He envisioned a Westernized Indian society. He aligned with the views of Charles Grant and believed that the traditional Eastern model of education was insufficient. He placed great emphasis on European and Western literature and also promoted the religious dimension of learning by spreading Christian beliefs and philosophy. Being loyal to his country, he worked actively to propagate British culture and the English language ([Sharma and Nemade, 2018](#)). Lord Hardinge suggested that administrative positions in the Company should be given to those with an English education. As a result, English studies soon became essential for employment, prompting people to learn the language to secure work ([Sharma and Nemade, 2018](#)). The first official report on secondary education, often called the Magna Carta, marked a turning point. It increased the government's responsibility and accountability in developing secondary education in India. In this way, universities were established in Bombay, Madras, and Calcutta. Additionally, many educational institutions began receiving financial assistance under what came to be known as the “grants-in-aid” system ([Bhattacharya and Sachdev, 2021](#)).

After gaining independence, India quickly started formulating its education policy with the aim of providing free and compulsory education. Several constitutional provisions allowed citizens access to education and the freedom to pursue learning at institutions of their choice, including opportunities in religious and women's studies. After independence, both the central and state governments placed strong emphasis on strengthening secondary education, focusing on its expansion and improvement. This was considered a key driver for national development and social transformation. Several commissions and initiatives were undertaken to implement reforms in the education sector of India after independence.

The key reforms are as follows:

1949 University Education Commission: - The first commission formed after independence was the university commission, was known as the Dr Radha Krishna Education Commission. It focused on Higher Education and reforms in curriculum.

1952 Secondary Education Commission: - The Mudaliyar Commission, chaired by Dr Lakshana Swamy Mudaliyar in 1952, mainly focused on school education.

1966 Kothari Commission: - D. S Kothari was the chairman of the commission. The purpose of this commission was to identify the problems and reasons for public disagreement and dissatisfaction, and to outline a national policy.

In 1968 National Policy on Education marked India's first official educational policy, developed following the suggestions of the Kothari Commission. Its purpose was to establish a cohesive and modernized educational framework throughout the nation. The policy focused on raising educational standards at every level and guaranteed “free and compulsory schooling for all children up to 14 years of age in nearby institutions” ([Jeena and Sabu, 2021](#)).

The 1992 Programme of Action was a key initiative that advanced education reforms and ensured the practical execution of India's National Policy on Education. The Action Programme aimed to improve the quality of education, provide education to all, and reduce social disparities.

2009 Right to Education Act: - The 86th amendment to the Indian Constitution in 2002 added Article 21-A, which ensure Education made free and mandatory for children aged 6 to 14.

2012 Justice Verma Commission on Teacher Education: - The commission recommended the establishment of institutions to improve the quality of teacher training both before and after service. It also suggested conducting teacher audits, ensuring punctuality and attendance.

2020 National Education Policy: - It aims to create developmental reforms in education and brought holistic approach in teaching method.

With the advent of colonial education systems, education shifted towards a classroom- centered, standardized curriculum designed for mass instruction ([Kumar, 2005](#)). This system emphasized literacy, examinations, and vocational training, but it often overlooked the individualized learning aspect that Gurukuls had fostered. Education is important for framing and modernizing society. It determines the quality of human resources, which, in turn, influences the overall transformation of society. Elements of the ancient education system, especially its best practices, can be thoughtfully

integrated into the modern educational framework. Due to rapid technological advancements, significant improvement could be noticed in education system. It is essential for the younger generation to understand how to use this technology to promote knowledge and preserve cultural values. Visionaries like Rabindranath Tagore recognized the need for reform in India's education system early on. Today, many children spend their parents' hard-earned money on education, yet still struggle to receive quality instruction and face challenges in securing meaningful employment. This highlights the urgent need to align education with real-world skills and opportunities. In contrast, modern education focuses on the holistic development of an individual's personality. It uses technology to actively engage students and support them in developing their skills at their own pace and according to their learning needs. The primary goal of modern education is to nurture well-rounded individuals by offering comprehensive training and unlocking each student's hidden potential for the benefit of society.

During the colonial era, many schools and colleges were established by the British educational framework. Higher Education Institutions such as the "University of Calcutta (1857), the University of Bombay (1857), and the University of Madras (1857) were established to offer higher education" (Gautam, 2016). In the post-independence period, there was a substantial expansion in both primary and secondary education. The good initiative of the government has introduced significant transformations in India's modern education system. From time to time, the government has introduced useful technologies to make education more accessible. Efforts were directed toward enhancing infrastructure, teacher training, and curriculum reform. Educators already use AI-driven tools in their daily routines, including voice assistants at home and applications that help with grammar correction, sentence completion, and essay writing. Many educators are exploring and trying out newly released AI tools. The education sector bears the responsibility of leveraging AI to advance educational goals while also safeguarding against potential risks associated with its integration into educational technology. This AI systems are evolving rapidly and offer significant support to both teachers and learners. However, when AI functions behind the scenes such as in infrastructure it raises concerns about transparency and trust. Although AI can seem almost

magical in new applications, educators and procurement teams must demand evidence of effectiveness. AI remains a broad and complex area that needs more involvement of stakeholders ([Elazab, 2024](#)). There is a need for AI policies focused on education that uphold civil rights and democratic principles throughout the design, implementation, and oversight of automated systems nationwide. The use of AI tools in education must prioritize student safety and effectiveness. These systems should include protections against algorithmic bias, ensure data privacy, offer clear notifications and explanations, and provide human oversight and avenues for addressing problems when they occur. AI-powered educational systems are likely to be appealing, in part, because they can facilitate more natural and dynamic interactions in teaching and learning. Traditional educational technology platforms often restrict the ways teachers and students engage with the tools. For instance, users may be limited to selecting options from a menu, answering multiple-choice questions, typing short responses, dragging objects, or using touch gestures. With the integration of AI, students can explore, invent, and develop creative ideas through advanced models that support open-ended, innovative tasks ([Elazab, 2024](#)). However, this progress also raises an important question: could AI potentially limit students' critical thinking abilities?

The introduction of the NEP marks a critical step in assessing both the quality and reach of educational institutions across the nation. Interestingly, studies indicate that the ancient Gurukul system provided universal and free education without state involvement a concept that remains relevant even today. The idea of merging the traditional Gurukul model with modern technology has become an emerging area of interest in alternative education methods. Efforts to implement computer-aided Gurukul education are currently in a largely theoretical stage. Although cloud computing holds promise for delivering personalized learning experiences, practical outcomes are yet to be realized. Nevertheless, combining technological advancements with the core principles of the Gurukul system has the potential to enhance the quality of education in India and yield meaningful, positive results. By introducing modern technology with traditional methods helps in forming the flexible learning through virtual gurukuls, AI can support personal, technology for spirituality teach students to tackle stress and practice mental wellness and gurukul adopting sustainability use

eco-friendly materials, solar energy and reducing waste materials ([Bhattacharya and Sachdev, 2021](#)).

“The establishment of the Secondary Education Commission (1952–1953), the University Grants Commission (UGC) in 1956, and the Kothari Commission (1964–66) represented the modernization of India's education system” ([Dubey and Nimje, 2015](#)). Digital learning technology has proven to be a powerful tool for educators, helping to engage students and spark interest in learning. As India emerges as a global leader in the field of computing, it is important to recognize the country's rich legacy in this domain. The rise of India as a leading destination for offshore software development was fuelled by visionary policies that aimed to modernize the nation and uplift its people from poverty. These strategic initiatives laid the groundwork for India's emergence as an IT powerhouse. The Digital India initiative was launched with the creation of the National Informatics Centre. Since 1984, various State Governments have allowed private institutions to establish schools and colleges, greatly expanding educational access across the country. AICTE Act of 1987 was established to uphold the quality of technical education. This council is responsible for promoting excellence in technical education, ensuring planned and coordinated growth, and maintaining standard regulations. Contemporary education emphasizes academic fields, particularly STEM (science, technology, engineering, and mathematics), while also embracing subjects like the arts, humanities, and vocational studies. Its central aim is to equip learners with the skills needed for career success and meaningful contributions to society (Dolas and Mandaogade, 2025). The advancement in digital learning in education has significantly improved how information is delivered, making learning more engaging and effective. Global research shows that ICT enhances student participation and improves teaching-learning methods. Tools like smart boards, interactive projectors, and multimedia technologies have revolutionized classroom instruction. They have also facilitated the creation of educational resources such as videos, computer-based training modules (CBTs), and digital learning content. To make quality educational materials widely accessible, National Programme on Technology Enhanced Learning in 2003 was launched. This e-learning initiative focuses on university-level STEM subjects.

NEP 2020 highlights the importance of integrating AI with the education sector. It also promotes the development and use of both hardware and software based on advanced technologies such as AI, machine learning, big data, blockchain, smart boards, and adaptive systems. These tools are designed to improve students' learning experiences and customize educational pathways according to individual requirements. Furthermore, the policy supports the creation of AI-driven platforms to monitor and analyse students' academic journeys. By collecting and interpreting learning data, these systems can identify students' strengths, weaknesses, and interests, enabling them to make well-informed choices regarding their future career directions. Since its implementation, India has made significant progress in enhancing the efficiency and effectiveness of learning through AI-driven technologies ([Gowda, 2023](#)). Hybrid approaches, like combining ChatGPT with community-driven mentorship groups, can restore the Gurukul's focus on relationship-centered learning while tackling modern-day inequalities. For instance, India's NEP 2020 promotes integrating AI with instruction in regional languages. Evidence shows that these efforts have positively impacted the country's education system. AI has played a key role in helping India move toward achieving Sustainable Development Goal 4 (SDG 4), promoting greater equality, equity, and inclusivity in education. The thoughtful and responsible application of AI, with a focus on human-centered design, has advanced considerably in recent years. Digital technology advancements like Intelligent Automation, Machine Learning, Deep Reinforced Learning, Metaverse, 4D Printing, Cleantech, GenAI, Intelligent Learning, Immersive Learning, Personalised Learning, Intelligent Tutoring, Metaverse and deep learning has caused a major shift in various domains, including education, image and speech recognition, manufacturing, and automation. AI's rapid evolution has enabled its application across diverse sectors such as Education, agriculture, cybersecurity, healthcare, and industry ([Gowda, 2023](#)). In the context of education, AI can support policymakers by providing accurate insights and data-driven recommendations to inform reforms. UNESCO also shows the transformative potential of AI in education. As a result, numerous AI-based techniques and tools are now being explored and implemented to elevate educational experiences and outcomes ([Gowda, 2023](#)).

The introduction of AI-focused courses in education. The objective was to equip students with AI skills that align with current and future job market demands. In line with this, the All-India Council for Technical Education introduced policy approving AI-centric engineering programs, leading many engineering colleges to launch undergraduate courses in AI and related fields. Additionally, several schools began offering AI-related subjects to students at earlier educational stages. The policy underlines that India must lead in training professionals in emerging and rapidly advancing fields such as big data analytics, and genomic sciences especially given their relevance to health, environmental sustainability, and improved living standards. These subjects are intended to be embedded into undergraduate curricula to enhance youth employability ([Gowda, 2023](#)). To promote scientific thinking and technological progress, the government founded premier institutions such as IITs and IIMs in India's development. This period also witnessed the modernization of education along with the incorporation of technology. There is a significant shift noticed in education with the coming of e-learning online platforms. Digital classrooms and interactive learning in 21st century show transformation in education. Now digital classrooms have interactive whiteboards, projectors and tablets that makes students more engaged and focused learning environment. This helps the students to understand any topic with crystal clear. Advancements in technology have reshaped learning environments, making education more accessible, interactive, and personalized. Classrooms have evolved significantly from using traditional blackboards to incorporating digital learning tools. The emergence of the World Wide Web and computer-mediated communication marked the beginning of online and distance education, enabling students to access learning materials and resources from virtually anywhere. Today's technologies, including AI and mobile devices, play a crucial role in modern education. They provide instant feedback, support diverse learning styles, and enhance academic outcomes. These tools help educators present content more visually and interactively, greatly aiding student understanding and engagement. The use of digital resources such as eBooks and digital library makes easier to access the works of various authors. This concise history of educational technology highlights how even the earliest tools laid the groundwork for today's

advanced systems, ultimately enabling students to interact with learning materials more deeply and paving the way for computer-based instruction in classrooms.

But in today's times there are various ways to make understand any parts of subjects through easy method by using of digital technology like 3D images and video and summarizing the section with the help of Artificial intelligence. This reinforces the students learning and understanding the concept of subjects. With the help of AI work would become accessible and maximise the things quickly and minimize the error and risk. But in current times with the help of machine learning and AI, complicated diseases could also be treated quickly and made study by the medical students. Even the engineering of any buildings like roads, bridges become easier for the students to learn and make modification through the AI. Though they applied the traditional concept but the high-tech education in the modern education has made more accessible and concrete in constructing building. The expansion of internet access allowed learners to go beyond traditional textbooks, giving them the ability to explore information from around the world. Concepts like online education and blended learning have redefined traditional academic models by offering more flexible and accessible learning experiences. AI-powered intelligent tutoring systems have further personalized education by addressing individual learning gaps, making the process more effective. AI stands out as one of the most transformative developments in educational technology, providing innovative tools that broaden teaching methods and significantly enhance learning outcomes.

4 CRITICAL PERSPECTIVE ON AI IN EDUCATION

AI in education has been challenging the technical development and educational innovation. With evolving time, AI has significantly changed the educational landscape through efficiency and accessibility. One of the greatest challenges in AI-driven technology is the digital divide between the urban and rural areas due to lack of basic availability of economic resources and infrastructure. However, rapid development in technological sectors brought major problems and challenges in the society as well. In many

places especially in rural and countryside areas, students are marginalised from using digital technology due to lack of availability of the resources. There should be shift to be made to include AI role in addressing current inequalities regarding access to knowledge, research and the diversity of cultural expressions and to ensure AI does not widen the technological divides. It should be accessible by everyone and can take advantage of the technological revolution under way and access its fruits, notably in terms of innovation and knowledge. In the contemporary times big tech companies such as Amazon, Google, Microsoft and Salesforces have been involving and investing in education infrastructure through the AI as a service system. As for example, Amazon has declared that AWS cloud would help schools, universities and other educational institutions to analyse and monitor the student's performance and the deployment of machine learning would help in predicting student outcome and thus AI could enable teaching assistants and tutors which helps in developing personalized learning outcomes. This companies stores data in large number which raises multiple risks and challenges regarding the data privacy and protection to individual data stored in the educational institutions and organisations. This example shows that how digital companies are in competition for dominance over the digital infrastructure of education. Hence, there is a need for policy debates and regulatory frameworks to be made to avoid misuse of technology. Therefore, it is important to establish national and state level policies to ensure inclusive participation of all sections of society and also particularly for under-resourced schools. Strong frameworks are required for ethical governance and transparency. As education becomes more technology oriented it requires ethical based structures and frameworks. Therefore, future education-based learning, not only required traditional method and digital method but also approaches ethical based understanding to serves human development rather than undermining it.

5 CONCLUSION

This chapter discuss about the wider historical developments and transition from traditional education system of ancient times in Gurukul to Modern

times in digital technology like AI contexts. It also discussed about the reforms taken by government in modernisation of Indian education system. It also highlights the need for AI education to enable India to become a 'digital superpower'. Digital technology has made the teaching and learning process more flexible. In this way learners would be accessible to learning through changing patterns of social and economic needs. This can promote skill renewal and intellectual growth and progress among the individuals. AI and digital learning transformed the modern teaching and learning method, which has become seminal. This created the path of progress and development for the future learning through different forms and establishing dynamics reforms in educations.

The emerging digital technology provides highly skilled forms of instruction, reshaping the traditional methods of teaching and learning. The Indian Knowledge Systems represent transformative changes which reinforces India's ancient intellectual traditions and positions them as important and practical tools for achieving holistic development and growth in modern times. With development of AI and advancement in IT sectors, the changes in the Indian society could be noticed in greater forms. This brought a major improvement in Education, Health, Automobiles etc. The workload of the students and teachers become less in the academic world and saves much time for example in taking attendance, like chatbots check the assignment, performance and grade the marks of the students. NEP 2020 brought significant changes in the curriculum of school syllabus and reforms in the academic works. Integration of IKS with digital technology can bring positive outcome and produces effective in students' spiritual, intellectual, emotional and physical growth, shaping individuals who are connected to their traditional and cultural identity. India can make become self-reliance and enlightened nation in all aspects. But with the addition of Gurukul system and Digital technology together would bring more inclusivity and modernise the education system. This could be possible through the government initiative and policies by taking out strong initiatives and reforms.

REFERENCES

- Adhikari, T.N., (2025). Inclusiveness in Gurukul education system. *Gyanjyoti: International Multidisciplinary Journal for Advanced Research*, 3(1). Available at: <https://doi.org/10.3126/gyanjyoti.v3i1.53040>
- AWS (Amazon Web Services). (2020). Machine Learning in Education. Amazon. <https://aws.amazon.com/education/ml-in-education/>
- Bhattacharya, S., and Sachdev, B.K. (2021). Gurukul system versus modern education in India – a need for amalgamation of the two systems to eliminate the crisis of illiteracy, economy and social problems of the society. *International Journal of Recent Advances in Multidisciplinary Topics*, 2(11), November, pp. 102–104. [↵](#)
- Biswas, A.K., (2016). Development of education in India during the medieval period: a historical approach. *International Journal of Research and Analytical Reviews (IJRAR)*, 3(2), pp. 87–91. [↵](#)
- Dharampal, P., (1983). *The Beautiful Tree: Indigenous Indian Education in the Eighteenth Century*. New Delhi: Biblia Impex, p. 272–278
- Dolas, P.V., Mandaogade, P. and Bajaj, P.B., (2025). A comparative study on Gurukul system and modern educational system., 20(1), pp.299–303.
- Dubey, K.T., and Nimje, A.A. (2015). A study of critical comparative analysis of ancient Indian education and present education system. *International Journal of Advance Research in Science and Engineering (IJARSE)*, 4(Special Issue 01), March, pp. 814–815. [↵](#)
- Elazab, M., (2024). Artificial intelligence and the future of teaching and learning. *Artificial Intelligence Arab Forum*, 23(2), pp. 1–16. [↵](#)
- Gautam, M., Singh, S., Fartyal, G., Tiwari, A. and Arya, K.S., (2016). Education system in modern India. *International Journal of Scientific Research and Education*, 4(1), pp. 4876–4879.
- Gopalkrishnan, S., (2023). New Education Policy 2020 in India: future rewinds to the past. *International Journal of Inclusive Education*, pp.1–15. [↵](#)
- Government of India. (2020). National Education Policy 2020: <https://nenow.in/wp-content/uploads/2020/07/NEP-final-for-circulation.pdf> [↵](#)
- Gowda, R.V.M., (2023). Education 5.0: Evolution of promising digital technologies – a comprehensive review. *International Journal of Advanced Science and Engineering*, 10(2), pp. 3422–3448. [↵](#)

- Jain Shelly, K., (2015). Relevance of Gurukul education system in present circumstance: A philosophical perception. *Journal of Philosophy, Culture and Religion*
- Jeena, P.M. and Sabu, P.J., (2021). Role of committees and commissions on the evolution of educational policies in India. *NIU International Journal of Human Rights*, 8, p. 4 [↵](#)
- Joshi, D., (2020). Gurukul and modern education system in India: Holistic outlook. [↵](#)
- Kumar, K. (2005). *Political Agenda of Education: A Study of Colonialist and Nationalist Ideas*. Sage Publications. [↵](#)
- Madhekar, M., (2020). Perception of integration of Gurukul system in modern Indian education: A study into quality education. *International Journal of Research and Analytical Reviews (IJRAR)*, 7(3) [↵](#)
- Mookerji, R.K., (1989). *Ancient Indian Education: Brahmanical and Buddhist*. Delhi: Motilal Banarsidass, p.18.
- Oke, S.A., (2008). A literature review on artificial intelligence. *International Journal of Information and Management Sciences*, 19(4), pp. 535–570.
- Pareek, A.P., (2021). A survey on Gurukul education system. *Journal of Contemporary Issues in Business and Government*, 27(3). Available at: <https://cibg.org.au/> [↵](#)
- Pathania, A., (2011). Teachers' role in quality enhancement and value education. *Academe*, 24(1), pp. 19–25.
- Selvamani, P., (2019). Gurukul system – an ancient educational system of India. *International Journal of Applied Social Science*, 6(6), pp. 1620–1622. [↵](#)
- Shanwal, V.K., (2023). Development of the Gurukula education system in India. *Journal of Education and Teacher Training Innovation*, 1(2), pp. 60–67. [↵](#)
- Sharma, R. and Nemade, P.N., (2018). Changing faces of education system from ancient India to present digital education in India. *IOSR Journal of Engineering (IOSR-JEN)*, 8, pp. 6–11. [↵](#)
- Vyāsa, (1947–1966). *The Mahābhārata: Śāntiparvan*. Critical Edition, ed. Sukthankar, V.S. et al. Poona: Bhandarkar Oriental Research Institute, 34. 13.

Zaheer, S. and Jafri, H., (2012). Education and transmission of knowledge in medieval India. *Intellectual Discourse*, 20(1), pp. 79–102.

Author Index

Aggarwal, G. [146](#)

Ananya [157](#)

Arslan, A.C. [74](#)

Arya, A K [83](#)

Bakhodir Bakir Ugli, E. [95](#)

Bhardwaj, R. [83](#)

Bula, Germán [31](#)

Chakravarty, U. [60](#)

Chatnani, N.N. [39](#)

Chhabra, L. [39](#)

Dan, I. [187](#)

Gera, G. [103](#)

Ghosh, P. [83](#)

Gupta, A. [103](#)

Kaur, M. [178](#)

Kaymaz, Y. [74](#)

Krishna, K.R. [52](#)

Kumar, A. [178](#)

Kumar, R. [178](#)

Kumar, S. [128](#), [137](#)

Kumari, S. [67](#)

Kumar Rath, S. [103](#)
Kumar Shanu, S. [119](#)

Lawrence, S. [8](#)

Maseili, M. [52](#)
Mehta, L. [103](#)

Nair, A. [83](#)

Pandey, A. [103](#)
Priyadarshi, A. [11](#), [83](#)

Rai, N. [165](#)
Rai, P. [137](#)
Raj, A. [157](#)
Raj, V. [187](#)
Ranjan, A. [146](#), [175](#)
Rath, A.R. [11](#)
Rath, S.K. [11](#)

Sachdeva, C. [103](#)
Salil, S. [60](#)
Sharma, S. [103](#)
Shrinag, P. [83](#)
Shrivastava, S. [103](#)
Singh, A. [119](#)
Singh, D. [1](#), [5](#)
Singh, K. [114](#)
Singh, N. [95](#)
Singh, R. [83](#)
Singh, S. [128](#)

Tashtemirovich, A.O. [178](#)
Thakur, S. [83](#)
Tiwari, V. [103](#)

Uktam Panjiyevich, J. [95](#)