

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

Deptartment of Humanities and Social Sciences

Artificial Intelligence and the Human Condition

Symposiasts

Professor Karamjit Gill,
Professor T.V. Prabhakar
Professor (Rtd.) Rahul De
Professor (Rtd.) Y.N. Mohapatra
Dr. Anirban Bandyopadhyay
Dr Nikhil Mahant

Agentic* AI ⋆ AI Ethics ⋆ Artificial General *Intelligence* ⋆Human Agency ⋆*Algorithmic* Bias ⋆*Epistemic* Injustice ⋆AI Governance ⋆Meaningful Work ⋆ *Post-Work

4th Floor,
DOAA Conference Room,
FB Annexe Building

26/11/2026 09:30

SCHEDULE

9:30 - 9:45
Inauguration

9:45 - 10:00 am
Professor Vineet Sahu
Introduction of the theme of the symposium

10:00 - 10:45 am
Professor Rahul De
The Ethical Challenges Posed by the AI Industry

10:45 - 11:30 am
Professor Karamjit Gill
AI Trajectory—from ELIZA to Agentic AI

11:30 - 11:45
Tea break

11:45 - 12:30
Professor Prabhakar T. V.
AI - Nemesis or Opportunity?

12:30 - 1:15 pm
Dr. Anirban Bandyopadhyay
Beyond Logic: AI, Human Agency, and the Responsibility to Judge (online)

1:15 - 2:30 pm
Lunch break

2:30 - 3:15 pm
Dr. Nikhil Mahant
Dr. Nemi Pelgrom
Method Selection, Will, and General Intelligence

3:15 - 4:00 pm
Professor Y. N. Mohapatra
The Incommensurable Gap: Remaining Human in an Age of AI

4:00 - 4:30 pm
Tea break

4:30 - 6:00 pm
Panel Discussion

MODERATOR
Dr. Sushruth Ravish

PANELISTS
Ms. Jyoti Kishore
Dr. Kshitish Sharma
Mr. Sanjiv Sikri

6:00 - 6:15 pm
Dr. Sushruth Ravish
Concluding Note
Vote of Thanks

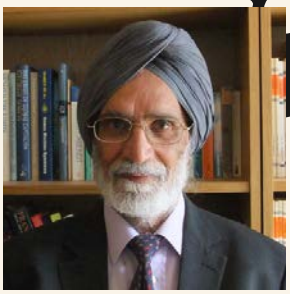
THE ETHICAL CHALLENGES POSED BY THE AI INDUSTRY



Professor Rahul De
Founder & CEO, Memoric AI
Professor (Rtd.) IIM Bangalore

The emerging AI industry consists of many players. At the base are the firms that make the chips, firms that provide energy for the data centers, those who collect and process data, and also those who train the AI engineers. Then there are the foundation model labs, mainly in the USA and also in China, Europe, and India, who create these millions of large and small models. Then there is the whole set of infrastructure providers who host these AI models. The tool and application builders create the tools that we use. The consuming end of the industry are the users: organizations, firms, governments, and society at large. This talk focuses on the ethical challenges posed by various players in the AI industry, consisting of both producers and consumers of AI products. We will consider cases of AI data centers, data labeling, production of copyright-protected content, and aspects of cognitive harm.

AI TRAJECTORY-FROM ELIZA TO AGENTIC AI



Professor Karamjit S Gill
EIC, AI & Society,
Professor Emeritus,
University of Brighton

Given the techno-utopian narrative that our future lies with Generative AI, especially Agentic AI, Then why are we now facing an ethical dilemma of embedding AI in societies?

In other words, what is it about the engineering of Agentic AI that raises ethical concerns about its impact and implication for societies. For example, how should universities integrate AI in educational processes and mechanisms, how do we cope with the implications of AI for social, cultural, economic, governance landscapes. The talk explores these questions and asks: how do we mould socially and culturally responsible AI for societal benefits whilst mitigating the danger of falling into the trap of Faustian seduction in which we bargain our soul with the machine in exchange for getting what we desire. During the 1980s the challenge and concern was how to mitigate the consequence of turning 'judgment' to 'calculation', during 1990s and early 2020s the concern was about the impact and implication of turning the human (outer-inner) into data.

Now the challenge and concern are: what we mean by interaction with Agentic AI and this raises the question of what we mean by concepts such as ethics, ethical/moral machine, ethical governance. Some of the key issues we face are: How do we design AI tools that align with societal goals such as those of trust and trustworthiness? Can the idea of the ethical machine drive towards taming ethical and moral dilemmas of alignment, beyond the Cartesian belief in human-AI co-evolution. How to tame the rise of “Destiny Machines” (e.g. social media platforms, Agentic AI) from becoming the contemporary manifestation of human destiny, relinquishing our ability to shape the human future by placing our faith in external forces. By going along with this faith, are we limiting the human potential whilst avoiding the responsibility of determining the destiny of humanity? By reflecting on these issues, we may ponder how the convergence of concepts of place, individualism and second cybernetics became a catalyst in the evolution of AI into the Black Box of Agentic AI.

In exploring these questions, we ask whether present day proponents of the Cartesian faith allow themselves to heed Weizenbaum's warning that humans and computers belong to separate and incommensurable realms.

ARTIFICIAL INTELLIGENCE - NEMESIS OR OPPORTUNITY?



Professor T V Prabhakar
Distinguished Visiting Professor,
WSAIS, IIT Kanpur

The accelerating development in AI capability clearly has the potential to disrupt society. Historically, all technological innovations have done this, while eventually improving the quality of human life. However, with AI this disruption appears to be at an unprecedented scale.Jobs are being lost, certainly being reduced. All labor either physical, intellectual or even creative is being encroached. We believe this is a pivotal moment for humanity, forcing us to re-invent ourselves and explore the deeper meaning of life and existence. In this talk we will discuss how AI will impact our lives in the short term. What the technology has to offer and how we can use it. We will talk about a fully developed AI, often called AGI, and how it is different from a human and how that gives us an opportunity to redefine our lives.

BEYOND LOGIC: AI, HUMAN AGENCY, AND THE RESPONSIBILITY TO JUDGE



Dr. Anirban Bandyopadhyay
Principal Research Scientist
National Institute for Material Sciences, Japan

As artificial intelligence becomes increasingly capable of performing tasks that can be formalized, optimized, and evaluated according to explicit criteria, an important question arises: what happens to human agency when logical execution is no longer primarily a human responsibility?

I argue that the significance of the human role should not be sought simply in cognitive abilities that machines have not yet reproduced. Rather, the emerging human–AI relationship may force us to distinguish reasoning from judgment, competence from agency, and decision from responsibility. Many consequential human situations are underdetermined: objectives conflict, information is incomplete, contexts change, and the relevant values themselves may be uncertain. In such situations, intuition, contextual perception, improvisation, and the capacity to revise the framing of a problem become especially important. AI can participate powerfully in generating possibilities, analysing alternatives, and executing specified objectives.

But the delegation of these functions does not automatically delegate moral responsibility. Human agency may therefore increasingly reside not in performing every step of a reasoning process, but in determining which purposes should be pursued, which constraints should govern them, when established rules should be reconsidered, and who remains answerable for the consequences.

This perspective also bears on debates concerning the moral standing of AI. Increasing cognitive sophistication may compel serious philosophical examination, but behavioural competence alone does not settle questions of experience, personhood, or moral status. The deeper challenge of AI may thus be not the disappearance of humanity, but a sharper demand that we understand what human judgment and responsibility actually mean.

METHOD SELECTION, WILL, AND GENERAL INTELLIGENCE



Dr. Nikhil Mahant
Snr. Assistant Professor,
University of Western Australia (India Campus)
Dr. Nemi Pelgrom
Industry PhD Scholar,
Fortnox

This paper asks whether artificial entities – defined as algorithm-driven, autonomous systems – can achieve general intelligence, understood as the ability to perform a wide range of tasks across diverse contexts at or above human standards. Drawing on the distinction between intuitive 'System 1' and deliberate 'System 2' thinking in humans, we compare these ways of thinking in humans to probability-based (P) and logic-based (L) architectures in AI. Artificial entities based on P- and L-architectures both exhibit limitations: P-systems can at best simulate deliberate reasoning, while L-systems lack generality. Neurosymbolic approaches that combine P and L elements are thought to offer promise. This article identifies an ability that is necessary for general intelligence in artificial entities – the 'Method Selection Ability' (MSA) – and provides a formal treatment of the ability. MSA is the capacity to choose an appropriate strategy from a set of tasks, contexts, pipelines, and valuation functions.

We argue that each of these choices requires a standard of what matters, which no further method can supply without regress. In humans, such standards come from the will — the goals, desires, and sense of identity that make some outcomes matter to an agent. Current AI systems borrow these standards from their designers and users, but an artificial entity with something akin to a will of its own could achieve general intelligence.

THE INCOMMENSURABLE GAP: REMAINING HUMAN IN AN AGE OF AI



Professor Y N Mohapatra
Professor (Rtd.),
IIT Kanpur

The rapidly growing power and reach of AI based on large language models (LLMs) raise the spectre of a new incarnation of ‘Bhashmasura’. By challenging our understanding of human capacities, AI forces us to ask: ‘What remains that is distinctly me?’ Is there a gap between intelligence and being human that no amount of computing can close? Drawing on earlier proposals, I will explore possible frameworks for identifying this incommensurable gap—if it exists. I will also reflect on my struggles to preserve my agency in everyday use of AI and consider whether AI might help us become more human.

PANEL DISCUSSION

MODERATOR

DR. SUSHRUTH RAVISH

Assistant Professor,
IIT Kanpur

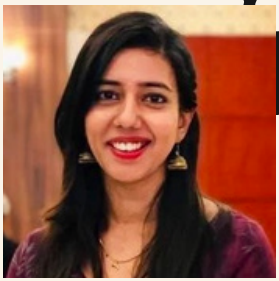
PANELISTS

MS. JYOTI KISHORE

DR. KSHITISH SHARMA

MR. SANJEEV SIKRI

ALGORITHMIC BIAS IN MACHINE LEARNING: SOCIETAL AND HEALTHCARE DISRUPTION BEYOND CLINICAL PRACTICE



Ms. Jyoti Kishore
PhD Scholar, IIT Kanpur

As artificial intelligence becomes deeply embedded across the healthcare domain, significant ethical anxieties emerge regarding biased algorithmic models. Existing scholarship highlights clear empirical harms, including systemic misdiagnosis, disparate treatment, and the neglect of vulnerable patient demographics. Yet, the broader societal and systemic ramifications of skewed clinical choices remain far more nuanced and subtle.

To truly evaluate these technologies, we must scrutinize automated decision-making past the immediate clinical environment, given that structural health disparities primarily originate in everyday social landscapes. Although clinical interventions represent only one vector influencing these divides, examining how medical disparities reinforce broader social inequities is essential to grasp the full toll of algorithmic bias.

This presentation investigates four key pathways whereby biased AI systems propagate far-reaching societal harm: the reinforcement of structural violence, the emergence of algorithmic paternalism, the deepening of epistemic injustice, and the degradation of institutional trust.

JAḌA PRAKṚTI, CETANA PUṚUṢA: AI AS INSTRUMENT, HUMAṆS AS AGENTS



Mr. Kshitish Sharma
Independent Scholar,
PhD, IIT (ISM) Dhanbad

Debates on the moral standing of artificial intelligence often assume that sophisticated cognitive performance signals consciousness. I question that assumption. Consciousness is defined here by the presence of prāṇa (life force), marked by six changes (birth, growth, reproduction, maintenance, dwindling and death) and three symptoms: thinking (jñāna), feeling (bhāva) and willing (icchā). By these criteria, AI is jaḍa (inert): it processes, but it does not live, feel or will. Drawing on Vedānta, AI is read as akin to prakṛti, inert yet self-functioning, and thus an instrumental cause (karaṇa), while humans remain the efficient cause (nimitta-kāraṇa) and the true agents (kartā). I can recall the 2007–10 comparisons of Google with God as a reminder that each wave of technological wonder tempts us to credit our instruments with what belongs to ourselves.

Yet agency has a sobering corollary: once an action is performed, our control over its consequences is limited, like a bomb already dropped.

Once the AI floodgates open, they cannot be closed. The pressing question is therefore not what we owe to AI, but what we owe one another. The answer lies in dharma, a moral discipline that centuries of pursuing “growth and advancement” have allowed us to evade.

DECIPHERING MEANINGFUL ACTIVITIES IN THE POST-WORK AI AGE



Mr. Sanjeev Sikri
Assistant Professor,
Wilson College

What if AI got it all right? It solved all the problems and enabled civilisation to achieve a deep utopia and absolute unemployment, sans anarchy. Machines have automated labour to a considerable extent; AI as a general-purpose 'thinking' tool now aims to automate core aspects of human cognition, judgment, imagination and creativity. To the extent that AI plans and efficiently executes several meaningful and meaningless tasks of an individual's life. In the post-work world, people may suffer from meaning deprivation. In such a world, I contend that man shall be bound to reinvent the notion of meaning, dissociated from work-related activities. With a narrow focus on meaningful activities that cannot be replaced by AI.