



VIGEYE VANI

Special Issue
on
**Artificial Intelligence and
Emerging Technologies for Ethical Governance**



Volume-LIV

January 2026



CONTENTS

Message from The Commission	03
Message from The Secretary	05
From the Editor’s Desk	07
Workshop Brief	09
Session-I: AI and Emerging Technologies: Past, Present and Future	13
National Policy Framework: National Strategy & Policy Initiatives including Global Perspective and Way Forward	14
Evolution of Artificial Intelligence & Emerging Technologies	17
Responsible & Ethical AI for Governance	20
AI for Enhanced Infrastructure Delivery, Transparency & Accountability	22
Predictive Cybercrime Analytics	27
Digital Tools in Preventive Vigilance and Good Governance at AIIMS New Delhi	30
Transforming Public Service Delivery: Fraud Detection in Government Health Schemes	33
Beyond Algorithms: India’s Experts Reflect on Trust, Privacy & the Future of AI Systems	37
Session-II: Panel Discussion with Experts from Industries	40
Leveraging AI & Emerging Technologies for Operational Efficiency & Preventive Vigilance	41



MESSAGE FROM THE COMMISSION



The Central Vigilance Commission observes Vigilance Awareness Week every year during the week of Sardar Vallabhbhai Patel's birth anniversary, with the focus on building awareness and reaffirming the commitment of everyone for upholding integrity. Vigilance Awareness Week 2025 was observed from 27th October to 2nd November 2025 on the theme:

“सतर्कता: हमारी साझा ज़िम्मेदारी”
“Vigilance: Our Shared Responsibility”

A notable highlight of this year's observance was a Workshop on **Artificial Intelligence and Emerging Technologies for Ethical Governance**, held on 31st October 2025. Bringing together Chief Vigilance Officers, Chief Technology Officers, senior officials, technologists, researchers and innovators, the Workshop provided a platform to reflect on the transformation underway in the vigilance ecosystem. Participants highlighted the vast potential of AI and machine-learning systems to analyze large datasets, identify anomalies, flag emerging risks, strengthen oversight in procurement and financial processes, and support real-time monitoring and decision-making. Equally, the deliberations underscored the ethical dimensions of AI - ensuring fairness, safeguarding privacy, preventing misuse, and maintaining accountability in automated processes - affirming that AI and emerging technologies must remain anchored in human values. As discussions unfolded through the sessions, the focus remained on how these technologies can evolve from being mere tools to becoming strategic allies in the fight against corruption. Thus, vigilance which has been predominantly punitive and has evolved to being more preventive and participative, may now evolve into being predictive too.

This special edition of *Vigeye Vani* presents the key insights from the workshop, along with articles and reflections contributed by the distinguished experts who participated in the sessions. Their perspectives illuminate how responsibly deployed digital tools can elevate vigilance from a reactive function to an anticipatory framework that prevents irregularities before they occur. The Commission encourages all organizations to integrate these learnings into their vigilance systems and continue fostering a culture where integrity is collectively upheld. By embracing innovation with responsibility and technology with ethical intent, public trust can be strengthened and governance systems made more transparent and accountable.

The Workshop marked an important step in equipping vigilance professionals with the knowledge to understand and adopt AI for ethical governance. The message was clear: The future of vigilance lies in harnessing the power of AI/ML responsibly, ethically and intelligently. When governed by strong values, these tools can transform the landscape of public integrity-building systems that are transparent, accountable and ahead of malfeasance.



(A.S. Rajeev)
Vigilance Commissioner



(Praveen Kumar Srivastava)
Central Vigilance Commissioner





MESSAGE FROM THE SECRETARY



igilance Awareness Week, which is observed annually to commemorate Shri Sardar Vallabhbhai Patel's birth anniversary, reaffirms our collective commitment to integrity and ethical governance. This year, the Vigilance Awareness Week 2025 was observed from 27th October to 2nd November 2025, alongside a preventive vigilance campaign from 18th August to 17th November 2025. This campaign focused on strengthening institutional processes through timely disposal of complaints, capacity-building initiatives, improvements in asset management systems and the adoption of digital tools for promoting transparency.

Across the country, organizations undertook a wide range of outreach programmes, awareness sessions, system studies and process reviews - ensuring that vigilance became a part of day-to-day governance rather than a symbolic observance. Several organizations, made concerted efforts embodying this year's theme, "**Vigilance: Our Shared Responsibility**".

The Chief Vigilance Officers played a key role in guiding these efforts, ensuring the structured implementation of the campaign, reviewing systemic vulnerabilities, and promoting digital tools for improved oversight. Their leadership has been instrumental in translating vigilance priorities into meaningful action on the ground.

A key highlight of this year's activities was the workshop on **Artificial Intelligence and Emerging Technologies for Ethical Governance**, held on 31st October 2025. Senior officials, CVOs, CTOs and domain experts discussed how vigilance in India is evolving from a reactive approach to one that is increasingly preventive, participative and predictive. The deliberations emphasized both the potential of AI for anomaly detection and risk identification, and the necessity of ethical safeguards, accountability and responsible use of technology.

This special edition of *Vigeye Vani* presents the core insights from the workshop, along with contributions from participating experts. I encourage all vigilance units and administrative leaders to draw from these learnings, and help create a culture of integrity, transparency and accountability. I also convey my sincere appreciation to the entire team led by **Shri Ashok Kumar and Shri Kapil Deo Narayan**, CTEs at the Commission, for their dedicated efforts in compiling and publishing this edition of *Vigeye Vani*.

A handwritten signature in blue ink.

(Trishaljit Sethi)

Secretary





FROM THE EDITOR'S DESK

I am pleased to present this Special Edition of the Commission's newsletter "*Vigeye Vani*" which showcases the key insights and contributions from the Workshop on **Artificial Intelligence and Emerging Technologies for Ethical Governance**, held as part of Vigilance Awareness Week 2025. This edition highlights the deliberations of senior officials, CVOs, CTOs and domain experts, reflecting the growing importance of technology-driven, preventive and predictive vigilance mechanisms in strengthening governance.

I extend my sincere appreciation to State Bank of India, our Knowledge Partner for the workshop, for their valuable support in curating content, sharing sectoral expertise and enriching the discussions featured in this newsletter. Their collaboration has significantly enhanced both the technical depth and practical relevance of this edition.

I also convey my heartfelt thanks to all members of the editorial team and the rapporteurs for their diligent compilation of material, thoughtful contributions and sustained effort in shaping this publication. Their collective work has ensured that this edition captures the essence of the deliberations and presents it in a clear and meaningful manner. I acknowledge the contribution of Shri Himanshu Jain, CVO, SECL, for his active support in the finalization and timely publication of this edition of *Vigeye Vani*.

I thank all participating organizations, contributors and experts whose insights have enriched this special edition. I am confident that the content presented here will assist vigilance units and administrative leaders in integrating emerging technologies responsibly and further strengthening systems of integrity, transparency and accountable governance.

A handwritten signature in blue ink, appearing to read 'Ashok Kumar', is positioned above the name and title.

(Ashok Kumar)

CTE, CVC



EDITORIAL BOARD...



Shri Ashok Kumar
CTE, CVC, Nodal Officer



Shri Kapil Deo Narayan
CTE, CVC, Member



Ms. Roopal Prakash
Director, CVC, Member



Shri Himanshu Jain
CVO, SECL, Member



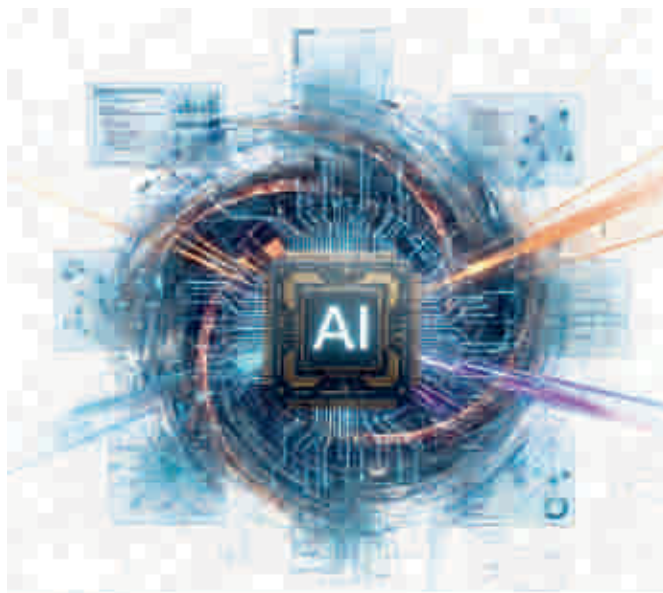
Shri Pradeep Kumar
CVO, IRCTC, Member



Shri Pradeep Khatana
TE, CVC, Member



Shri Prakhar Tripathi
TE, CVC, Member



Workshop conducted on 31st October 2025

Artificial Intelligence and Emerging Technologies for Ethical Governance

As part of the observance of Vigilance Awareness Week 2025, the Central Vigilance Commission organized a one day workshop on Artificial Intelligence and Emerging Technologies for Ethical Governance on 31st October 2025. The event brought together the Chief Vigilance Officers, Chief Technology Officers, senior officials from Central Ministries, PSUs, and leading experts from the technology and innovation ecosystem.

The workshop moderated by Shri Dhruvabrata Ghosh Dastidar, MD, Protiviti, began with contextual remarks on the transition of vigilance in India from its early punitive phase to preventive and participative approach, and now towards a predictive model supported by digital tools. Speakers emphasized that technological advancements are reshaping governance systems globally and that AI, with its capacity to process large datasets, detect patterns, and forecast risks, has significant potential to reinforce the integrity framework of public institutions. At the same time, there was consensus that technology must be applied responsibly, with safeguards to protect privacy and ensure fairness and accountability.

The workshop was organized in two sessions. Session I focused on the

evolution of AI, emerging technologies, and the way forward. Session II comprised a panel discussion on the application of AI for improving operational efficiency and preventive vigilance.

Session I: AI and Emerging Technologies – Past, Present and Future

Session I dealt with the evolution of AI globally and in India, capturing both doctrinal developments and practical applications across sectors.

Shri Abhishek Singh, IAS, Additional Secretary, MeitY, DG, NIC & CEO, India AI Mission, provided an overview of India's national AI strategy and its alignment with global standards. He highlighted how

The objective was to explore how emerging technologies-especially AI & ML can be integrated into the vigilance architecture to strengthen efficiency, transparency & ethical governance.





AI has moved from rule-based systems to machine learning, deep learning and generative AI and is now considered a critical tool for public service delivery, grievance redressal, fraud detection and regulatory compliance.

Prof. (Dr.) Anirban Dasgupta, IIT Gandhinagar discussed the technological foundations that enabled AI's rapid growth, including the advent of cloud computing, accelerators like Graphics Processing Units [GPUs] and Tensor Processing Units [TPUs], the breakthroughs in deep learning enabled by large training datasets, transformer models and the emergence of large language models. He underscored the importance of converging AI with Internet of Things [IoT], blockchain, and quantum computing to enhance transparency, traceability and optimization in governance.

Shri Deepak Goel, Scientist G & Group Coordinator MeitY, addressed the legal and ethical dimensions of AI. He elaborated on India's Digital Personal Data Protection Act and the importance of establishing frameworks for privacy protection, fairness and explainability. His remarks emphasized that trust in AI systems is essential for their credibility in governance.

Dr. Murali Jagannathan, Assistant Professor, IIT Madras presented case studies on AI-driven improvements in infrastructure delivery, including risk assessment of contract documents, benchmarking practices, enhanced correspondence management and the role of digital tools in improving accountability in large-scale public projects.

Shri Rajesh Kumar, IPS, CEO, Indian Cyber Crime Coordination Centre (I4C), highlighted the role of predictive analytics in combating cybercrime. He explained how AI tools automate evidence analysis, identify suspicious behavior patterns and forecast likely attack vectors, enabling law enforcement agencies to respond more proactively.

Dr. M. Srinivas, Director, AIIMS New Delhi, provided insights into the use of digital tools in preventive vigilance within the health sector. He elaborated on the use of technology for monitoring procurements, human resources, inventory and financial transactions.

Dr. Sunil Kumar Barnwal, IAS, CEO, National Health Authority, shared examples of AI's transformative impact on public service delivery, particularly in large-scale health systems.



The first session concluded with a moderated discussion and an open Q&A, where participants examined data quality challenges, algorithmic biases, infrastructural constraints and the importance of multi-stakeholder collaboration.

Session II: Leveraging AI and Emerging Technologies for Operational Efficiency and Preventive Vigilance

Session II brought together senior leaders from major sectors-including banking, payments, energy, manufacturing and services-to discuss real-world applications, practical challenges and strategic pathways for incorporating AI into preventive vigilance.

Shri Nitin Chugh, Deputy MD, SBI, emphasized how AI systems in the financial sector are being used for fraud detection, transaction monitoring, customer risk profiling and compliance automation. He noted that predictive vigilance tools have enabled early identification of

irregularities and enhanced oversight in branch-level operations.

Shri Nitin Mishra, Executive Director-Technology, NPCI, highlighted the enormous potential of AI in payment systems, where billions of transactions occur in real time. He discussed AI-based risk scoring, real-time anomaly detection and intelligent dispute resolution as critical innovations for safeguarding the digital payments ecosystem.

Shri Mohit Kapoor, Group CTO, Mahindra Group, spoke about the industrial transformation driven by automation, robotics, digital twins and edge analytics. He shared experiences on integrating ethics-by-design into enterprise AI systems and the importance of responsible innovation to prevent unintended consequences.

Shri Manish Malik, ED, IOCL, reflected on AI use in large-scale logistics, customer management and supply chain operations, stressing how data-driven vigilance can significantly reduce operational risks.

Shri Rahul Himalian, Director, IRCTC, shared experiences from the services and tourism sector, focusing on how AI-enabled monitoring has improved customer service and operational compliance.

Ms. Geeta Gurnani, CTO, IBM, emphasized global best practices in ethical AI, including principles of transparency, explainability, auditability and human oversight. She highlighted that organizations must invest in governance

structures and skilled workforces to successfully integrate AI into everyday operations.

The panel discussion highlighted that organizations using AI-driven tools have experienced substantial improvements-30-40% gains in operational efficiency and significant reductions in fraud-related losses. However, concerns about data privacy, algorithmic bias, inadequate infrastructure and the need for skilling remain key challenges.

Conclusion and Key Takeaways

The workshop concluded with reflections on how AI and emerging technologies can transition vigilance systems into a more anticipatory and intelligence-driven framework. Participants agreed that while AI provides powerful tools for risk detection and real-time monitoring, it must operate within a robust ethical and regulatory framework. Human judgment, institutional values and accountability mechanisms must remain central to its application.

The discussions across both sessions reinforced the message that technology and vigilance must work together-

technology offering precision, speed and analytical capability, and vigilance providing the ethical compass. The workshop successfully created a multi-stakeholder platform for knowledge-sharing and has set the stage for future capacity-building initiatives under the Commission.

This edition of *Vigeye Vani* captures the core deliberations, insights and key takeaways from the workshop, offering guidance to organizations as they work to integrate responsible AI into their vigilance and governance frameworks.



Session-I: AI and Emerging Technologies: Past, Present and Future



Sh. Abhishek Singh, IAS is an Additional Secretary to the Government of India and is presently posted as the Director General, National Informatics Centre with responsibilities of Artificial Intelligence & Emerging Technologies, Human Centered Computing (HCC) and Digital India Bhashini Division. He has a Master's degree in Public Administration from Harvard Kennedy School of Government. He is also an alumnus of IIT Kanpur.

Prof. (Dr.) Anirban Dasgupta is the N. Rama Rao Chair Professor of Computer Science & Engineering at IIT Gandhinagar. He did his PhD at Cornell University and subsequently worked at Yahoo Labs, Sunnyvale for 8 years before joining IIT Gandhinagar in 2014. His research interests include theory of ML, efficient ML and algorithms in general.



Sh. Deepak Goel is working with the Government of India as a Scientist G & Group Coordinator, MeitY. He has previously worked on the Data Protection and Online Gaming legislations at MeitY's Cyber Law Division. He holds a Bachelor's degree in Computer Science Engineering.

Dr. Murali Jagannathan is an Assistant Professor at Building Technology and Construction Management (BTCM) Division, Department of Civil Engineering, IIT Madras, with a specialization in construction contracts, claims management and construction dispute resolution. He holds a Master's degree from IIT Madras and a Doctorate degree from IIT Bombay.



Sh. Rajesh Kumar, IPS is working with the Government of India as the CEO, Indian Cyber Crime Coordination Centre (I4C), MHA. He has played a leading role in creating various digital platforms like Sahyog, Samanvaya, Suspect Registry and E-zero FIR. He is a graduate in Electrical Engineering from IIT Kharagpur.

Dr. M. Srinivas is the Director of AIIMS, Delhi. He is a Professor of Pediatric Surgery and is the former Dean of Employees' State Insurance Corporation (ESIC) Hospital and Medical College, Hyderabad. He has done his MBBS, MS, M.Ch [Pediatric Surgery] from AIIMS New Delhi. He completed his fellowship from Children's Research Centre, Our Lady's Hospital for Sick Children, Dublin, Ireland.



Dr. Sunil Kumar Barnwal, IAS is an Additional Secretary to the Government of India and is presently posted as the CEO, National Health Authority, Ministry of Health and Family Welfare. He holds Doctorate degree from ISM Dhanbad, a Master's degree in Public Policy from NUS & Harvard University and a B.Tech degree from ISM Dhanbad.

National Policy Framework: National Strategy & Policy Initiatives including Global Perspective and Way Forward



Abhishek Singh, IAS

Additional Secretary, MeitY
DG, NIC & CEO, India AI Mission



New Era of Governance

Artificial Intelligence (AI) has evolved from a futuristic concept to an indispensable force shaping the global order. From the early days of machine learning to the rise of deep learning and now to generative and agentic AI, the pace of change has been exponential. The “ChatGPT moment” of 2022 transformed AI from an expert tool into a household term. Today, millions interact daily with AI-drafting documents, analyzing data, or automating tasks-without realizing the extent to which it influences every layer of modern life.

Yet, for India, AI is more than a technological shift-it is a nation-building opportunity. The challenge and promise lie in channeling AI’s efficiency and intelligence toward strengthening public institutions, improving transparency, and ensuring ethical governance. The real measure of success will not be how many algorithms we build, but how effectively we use them to build trust, integrity, and inclusion in governance.

India’s AI Advantage

India enters the AI era with extraordinary strengths. Over the past decade, the nation has built one of the world’s most powerful Digital Public Infrastructures (DPI)-Aadhaar for identity, Unified Payments Interface [UPI] for payments, DigiLocker for paperless governance, and DIKSHA (Digital Infrastructure for Knowledge Sharing) for digital learning. Each of these platforms has demonstrated how inclusive technology can operate at population scale. The result is a fertile foundation for AI innovation. With more than 1.5 lakh startups, a dynamic talent base, and globally respected engineers leading efforts in OpenAI, Google, Meta, and Microsoft, India’s AI workforce is among the most capable in the world. According to the IndiaAI Expert Group Report (2023), the widespread application of AI could add USD 1.7 trillion to GDP by 2035 and create over four million new jobs. But beyond the numbers lies a deeper promise of governance that is transparent, data-driven, and citizen-centric.

AI for Ethical Governance

AI and emerging technologies hold the potential to revolutionize the ethics in governance. When applied with foresight, they can strengthen the three pillars of good governance-transparency, accountability, and inclusion.

- **Transparency & Accountability:** AI analytics can scan public expenditure data to identify anomalies, detect procurement irregularities, and flag inconsistencies in decision-making. Blockchain ensures traceable and tamper-proof records, while natural language tools make policies and circulars understandable to citizens in their own languages.

- **Fair and Evidence-Based Decision-Making:** Predictive models can simulate the outcomes of policies before rollout, helping eliminate bias and ensure equitable impact. Data-driven dashboards can track progress, highlight inequality, and recommend course corrections.
- **Citizen Engagement & Inclusion:** Multilingual AI chatbots, sentiment analysis tools, and participatory digital platforms allow citizens to provide feedback in real time, fostering a two-way dialogue between the state and the people.

As these technologies mature, they promise a new era of ethical governance—one where technology amplifies human judgment instead of replacing it.

Applications Transforming Governance

India's public sector is already seeing visible results from AI deployments across diverse domains:

1. Judiciary and Legal Research

The Legal Research Assistant (LEGRA) developed by NIC has transformed how the judiciary manages cases. Deployed across more than 2,500 courts, including near-universal adoption in Kerala and Karnataka, LEGRA provides real-time multilingual transcription and translation in 15+ Indian languages. Case processing time has reduced by up to 50%, enhancing both efficiency and accessibility.

2. Law Enforcement and Public Safety

AI-driven computer vision systems have enhanced crowd management and security during large events such as the Kumbh Mela. Predictive policing tools are being used in Bengaluru, Uttar Pradesh, and Punjab to identify crime hotspots and optimize resource deployment. Radio Frequency Identification [RFID]-based pilgrim tracking during the Amarnath Yatra has further demonstrated how technology can support human safety at scale.

3. Empowering Bharat through Bhashini

Language diversity has long been a barrier to digital inclusion. The Bhashini platform, a core component of the IndiaAI ecosystem, bridges this gap through AI-

driven translation, speech recognition, and voice synthesis across Indian languages. 26 lakh police officials now record multilingual FIRs and case diaries, 10 lakh government employees use voice-enabled eOffice noting and translation, 10.7 crore farmers receive real-time agricultural advisories in local dialects, 1.8 crore students and teachers benefit from multilingual AI learning assistants, and 50 crore banking customers access voice-based digital banking services. By enabling citizens to interact with government systems in their native languages, Bhashini is making governance truly participatory.

4. AI in Health, Agriculture and Climate

The IndiaAI Applications Development Initiative supports 30 prototypes addressing national priorities. These include AI-based early disease detection through enhanced X-rays, AI-powered vector-borne disease surveillance, precision agriculture for soil and water monitoring, and climate action models predicting disaster risk and resource allocation. Each initiative illustrates AI's potential for public welfare and sustainable development.

The IndiaAI Mission: Building the Future

The IndiaAI Mission, launched by the Government of India, is the strategic anchor for this transformation. Built around seven integrated pillars, it aims to create an inclusive, world-class AI ecosystem:

1. **Compute Capacity:** Over 38,000 GPUs accessible via 14 service providers, with subsidized use for startups and researchers at just ₹65 per GPU-hour.
2. **Foundation Models:** Development of indigenous large language models such as BharatGen (IIT Bombay) and Sarvam (Bengaluru), trained on Indian datasets and values.
3. **AIKosh Data Platform:** A national repository hosting 3,500+ datasets and 240+ models, enabling innovation through open access.
4. **Applications Development:** Sectoral challenges from health to climate and governance are being tackled through 30 selected AI solutions.

- 5. **Future Skills:** UG to PhD fellowships, 570 AI Labs through NIELIT (National Institute of Electronics and Information Technology), and NCVET (National Council for Vocational Education and Training) approved courses ensure skilling across disciplines.
- 6. **Safe and Trusted AI:** Ethical certification tools, bias detection frameworks, and the IndiaAI Safety Institute will ensure AI development remains transparent and accountable.
- 7. **Startup Financing:** Seed and global accelerator programs, including IndiaAI Station F in Paris, nurture an innovation pipeline from idea to impact.

Towards Responsible and Trustworthy AI

The conversation around AI must move beyond innovation to responsibility. India’s approach rests on three foundational principles: trust, transparency, and transformation. Current global models, trained primarily on Western data, often fail to represent India’s social and cultural diversity. Images of “an Indian man” or “an Indian home” in such systems tend to reinforce outdated stereotypes. To correct this, India is investing in contextual foundation models that reflect local values and realities. At the same time, the Government is advancing research in:

- Bias mitigation and fairness assurance to ensure that algorithms do not perpetuate discrimination.
- Privacy-preserving computation and data anonymization for secure digital ecosystems.
- Deepfake detection and watermarking to safeguard information integrity.
- AI explainability frameworks to make algorithmic decisions auditable and comprehensible.

These initiatives represent India’s moral commitment: innovation must always serve the public good.

The Road Ahead

The upcoming IndiaAI Impact Summit 2026 will showcase these achievements and foster collaboration between

government, academia, and industry. As AI systems become deeply embedded in governance, they will influence not only efficiency but also the ethical fabric of institutions. Used wisely, AI can detect anomalies in public finance, support vigilance processes, and improve accountability mechanisms—strengthening citizens’ trust in the state. The integration of AI, blockchain, and data analytics will make governance predictive, participatory, and preventive. When combined with platforms like Bhashini, it ensures every citizen—regardless of language or literacy—has equal access to services. This is the true spirit of Viksit Bharat: technology empowering people, not replacing them.

Conclusion

India stands on the threshold of an AI-powered governance revolution. The tools we build today will determine not just how efficiently we govern, but how ethically we lead. With its robust infrastructure, inclusive policies, and forward-looking mission, India is crafting a model where AI and ethics evolve together. In the years ahead, as AI takes its place beside electricity and the internet as a transformative force, India’s commitment to “AI for Good and for All” will define its leadership on the global stage. Through responsible innovation, collaborative institutions, and citizen-centric design, India is not merely preparing for the AI age—it is shaping it. With the right policies, partnerships, and purpose, AI can become the cornerstone of a governance system that is efficient, transparent, and above all, ethical—fulfilling the vision of a truly empowered and Viksit Bharat.



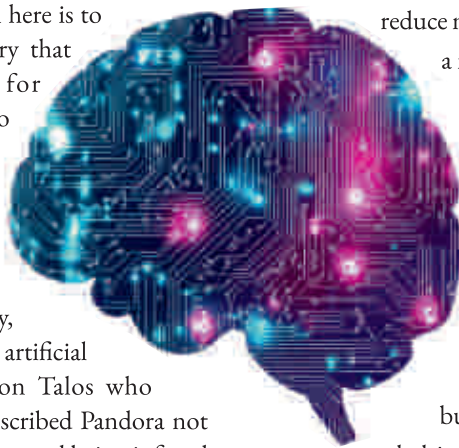
Evolution of Artificial Intelligence & Emerging Technologies



Prof. (Dr.) Anirban Dasgupta
IIT Gandhinagar

Artificial Intelligence is often described as a defining technology of our century, but its roots run far deeper than the innovations of the last few decades. To appreciate where we stand today, it is important to travel back through time and look at the ideas, myths, scientific breakthroughs and philosophical debates that shaped the modern field. The intention here is to narrate this evolution as a story—a story that humanity has been telling itself for thousands of years in its quest to understand and recreate intelligence.

Long before computer scientists ever wrote code or trained neural networks, early civilizations grappled with the idea of artificial beings. In Greek mythology, Hephaestus crafted mechanical servants, artificial lions and the giant bronze automaton Talos who guarded the island of Crete. Hesiod described Pandora not merely as a human but as an artificially created being infused with purpose. Across cultures, similar notions appeared—golems animated by sacred inscriptions, homunculi born from alchemy, and intricate automatons built by craftsmen such as Hero of Alexandria, al-Jazari and Jaquet-Droz. These creations show that the human imagination has always been fascinated by the possibility of intelligence made, not born.



The Enlightenment further advanced the idea through literature and early science fiction. Mary Shelley’s *Frankenstein* posed philosophical questions about creating life, while Karel Čapek’s *Rossum’s Universal Robots* introduced the very word “robot” to the world. Isaac Asimov’s stories then gave us the famed Three Laws of Robotics, framing ethical boundaries for machines long before such machines existed. Meanwhile, thinkers like Aristotle, Euclid, al-Khwarizmi, Ramon Llull and Gottfried Leibniz worked on formal reasoning, logic, algebra and computational ideas, unknowingly laying the foundations of modern AI. They believed that human reasoning could, at least in part, be reduced to mechanical processes—a radical idea at the time, but one that echoes directly in today’s algorithms.

The nineteenth and early twentieth centuries brought crucial developments. George Boole formalized logical operations, Alfred North Whitehead and Bertrand Russell attempted to reduce mathematics into a complete logical system, and David Hilbert challenged mathematicians to prove that all mathematical reasoning could be formalized. That dream collapsed when Kurt Gödel proved that no consistent formal system could capture all mathematical truths, a philosophical shock that also hinted at the limitations of intelligent machines. Alan Turing, building upon these ideas, demonstrated the halting problem and set theoretical limits on computability. Yet he also opened a door of extraordinary imagination by asking a deceptively simple question: “Can machines think?” His imitation game, later known as the Turing Test, became the earliest formal benchmark of machine intelligence.

The practical birth of AI began in the 1940s and 1950s with Warren McCulloch and Walter Pitts, who created the first model of neural networks. Marvin Minsky and Dean Edmonds simulated a rat navigating a maze using a network of vacuum tubes, and Arthur Samuel built a checkers-playing program capable of improving through experience, thereby coining the term “machine learning.” In 1956, John McCarthy convened the Dartmouth Conference, where the term “Artificial Intelligence” was officially introduced. Most of the giants of early AI—including Minsky, Newell, Simon, Shannon and others—attended. Their collaboration seeded the academic ancestry that would dominate the next half-century.

The years that followed saw intense enthusiasm. Systems like the Logic Theorist and the General Problem Solver attempted to mimic human reasoning. Rosenblatt’s perceptron demonstrated early learning capabilities. New programming languages such as LISP accelerated AI research. Robots like SHAKEY showed that machines could perceive their environment and perform tasks. Joseph Weizenbaum’s ELIZA demonstrated rudimentary conversational ability. But reality soon tempered expectations. Early machine translation failed because it lacked context, many problems proved too complex due to combinatorial explosion, and Minsky and Papert’s criticism of perceptrons halted neural network research for over a decade. Funding dried up, ushering in the first “AI winter.”

AI revived in the 1970s through expert systems, which encoded human expertise in rule-based logic. Systems like DENDRAL, MYCIN, INTERNIST and CADUCEUS demonstrated that AI could solve real medical and scientific problems, sometimes outperforming junior professionals. By the 1980s, corporations invested billions in AI, deploying expert systems widely. Japan’s Fifth Generation Project and renewed Western initiatives promised intelligent computers.



Yet oversized expectations again led to disappointment, and a second AI winter followed.

The late 1980s and early 1990s marked a turning point as AI embraced statistics, probability, and the scientific method. The introduction of Bayesian networks by Judea Pearl and others gave AI powerful tools to handle uncertainty. Reinforcement learning advanced through systems like TD-Gammon. Support Vector Machines emerged as a foundational technique. Robotics gained momentum with initiatives such as RoboCup. And in 1997, IBM’s Deep Blue famously defeated world chess champion Garry Kasparov, proving that machines could outperform humans in complex reasoning tasks.



The evolution of intelligent agents transformed AI once again. Autonomous systems such as NASA’s Mars rovers, iRobot’s Roomba, DARPA’s self-driving challenges, and later Google’s autonomous vehicles illustrated how algorithms, sensors, actuators and learning systems could work together to produce goal-directed behavior. Simultaneously, the explosion of digital data reshaped the field. Researchers realized that large datasets could sometimes outweigh algorithmic complexity, shifting focus toward data-driven learning. Breakthroughs soon followed: IBM Watson excelled at natural language question answering, virtual assistants like Siri and Alexa brought speech recognition to everyday devices, deep neural networks learned to identify images at unprecedented accuracy, and AlphaGo defeated the world’s best Go player—something once thought impossible.

All of this set the stage for the generative AI era. When tools like ChatGPT appeared, the world suddenly witnessed AI producing text, images and video with astonishing fluency. Yet generative AI is only the beginning. We are now moving toward agentic AI-systems capable of performing tasks autonomously, conducting analysis, drafting documents, assisting in diagnostics, reading balance sheets or performing customer service without explicit step-by-step instructions.

Beyond that lies physical AI, where robotics and machine learning merge to create machines capable of interacting with the physical world. And further still lies the research frontier of Artificial General Intelligence, raising profound questions about capability, responsibility, ethics and control. While the story of AI is full of remarkable achievements, it is equally important to absorb its lessons. Every wave of excitement has been followed by periods of disillusionment whenever expectations outpaced reality. This cycle reminds us that intelligence is complex, that limitations exist, and that

responsible development is essential. As AI rapidly enters governance, justice, education, healthcare and public administration, we must recognize both its enormous potential and its inherent risks. Biases embedded in models, incomplete datasets, hallucinations, misuse through deepfakes, and ethical ambiguity are challenges we must confront head-on. Despite these concerns, it is important to remain optimistic. For thousands of years, humans have sought to understand thinking, reason, perception and decision-making. Today, for the first time in history, we have the computational power, the scientific understanding and the engineering capability to build systems that can perceive, reason, predict and act in ways that augment human capacities. AI has progressed from mythology to machines, from philosophy to mathematics, from code to cognition, and now into a technology that influences nearly every aspect of our lives. The story is far from over, and much remains to be written.



The convergence of data, AI and ML is not just a technological shift but a strategic imperative to strengthen transparency, fairness and ethical governance.

As we move into an era where AI increasingly shapes public systems, governance and societal decision making, our responsibility is not merely to innovate, but to ensure that such innovation remains aligned with human values. If we continue to build cautiously, ethically and inclusively, AI will not only transform technology-it will deepen human understanding, expand opportunities and help us build systems that are transparent, efficient and fair. The long historical journey of AI reminds us that intelligence-whether natural or artificial-is best directed when it is guided by wisdom.

Responsible & Ethical AI for Governance



Deepak Goel

Scientist G & Group Coordinator, MeitY

Artificial Intelligence has transformed so rapidly that even the events of 2020 feel like a distant past. In the human world, five years is brief, but in the world of AI it is an entire era. As the gathering takes place during Vigilance Awareness Week, reflecting on transparency, accountability, and ethical governance, it becomes essential to examine not only how AI can support governance but also how its risks can undermine the very foundations of fairness and justice that public institutions strive to uphold. The concluding thoughts of the Central Vigilance Commissioner, emphasizing the principles of integrity, transparency, and responsible use of technology, provide an apt prelude to the concerns being highlighted regarding responsible and ethical AI.

A real example from a developed country illustrates how powerful these concerns are. In 2020, an AI-based fraud detection system known as SyRI (System Risk Indication) was deployed by Dutch Government for analyzing eligibility for welfare benefits. It processed details such as income, housing, education, ethnicity and neighbourhood crime rates to classify individuals by risk levels, supposedly to ensure that government benefits reached the right people. However, the system disproportionately flagged low-income immigrant families as high-risk, effectively excluding the very individuals for whom the welfare programme was created. The data used to train the model was undisclosed, the algorithms were opaque, and the entire system lacked transparency. Instead of

promoting fairness, the AI ended up discriminating against vulnerable communities. Public outrage followed, the democratic fabric was shaken and ultimately the system was declared unlawful for violating privacy rights and failing to distribute welfare equitably. This single case demonstrates how AI, when used without adequate oversight, can harm citizens and compromise the legitimacy of governance. A similar warning emerged closer to home when the Bombay High Court, in a judgment delivered recently, made a clear observation about AI-assisted faceless assessments in income tax proceedings. The Court held that when an authority exercises a quasi-judicial function, it cannot blindly rely on AI-generated outcomes. The findings produced by AI must be independently verified and cross-checked before forming the basis of any decision. AI may support analysis, but human judgment must remain paramount. The responsibility for fairness, accuracy and justice cannot be transferred to a machine.

These examples point to the need for oversight in all AI-supported governance. Oversight can take many forms, but the distinction between being in the loop and being on the loop is critical. In many administrative processes, we interact with AI systems that are capable of analysing data, summarizing documents and highlighting patterns far quicker than humans. Yet none of this sophistication replaces the essential role of human oversight. Systems that operate autonomously without meaningful human supervision risk becoming opaque and unpredictable. It is wiser, and safer, for officers to always remain on the loop, using AI for assistance while maintaining control over decisions and outcomes. The Digital Personal Data Protection Act, 2023 offers an important legal framework for trustworthy AI. It requires that personal data be collected with consent, used only for defined purposes, kept accurate, minimized to what is necessary and stored only for as long as needed. Even though government entities may process personal data without consent for lawful functions, the broader obligations of transparency, fairness and accountability remain firmly applicable. For AI, this means that the data used to train



models must be ethically sourced, its purpose clearly stated and its retention governed by defined policies. The DPDP Act thus serves not merely as a data protection law but as a governance toolkit that helps create trust infrastructure in public systems.

Fairness is at the heart of responsible AI because AI is not neutral. It learns from data that often carries historical biases, and in a diverse country like India, these biases can unintentionally target groups based on socio-economic factors, regional differences, language variations or other attributes. AI systems can exhibit what I call bias with confidence, where the model becomes increasingly certain about its skewed predictions and the human operator begins to trust those outputs without question. This is precisely why algorithmic fairness cannot be separated from India's constitutional ethos. Governance systems must ensure that AI does not disadvantage any community or individual, intentionally or unintentionally. As AI becomes a part of governance, vigilance must evolve as well. Traditionally, vigilance has focused on preventing corruption and ensuring ethical conduct, but in the AI era it must also include auditing the algorithm itself. Officers will need to understand how data is processed, whether models use data lawfully, whether adequate human oversight exists, whether individuals have

access to grievance redressal and whether the system treats all groups fairly. Vigilance is no longer confined to monitoring individuals; it must also monitor the digital tools that influence administrative decision-making.

The path ahead requires us to remember a simple but profound principle. AI can assist, accelerate and enhance our workflows, but it cannot replace human judgment. It can highlight patterns, but it cannot understand context. It can flag anomalies, but it cannot comprehend fairness. In the public sector, where decisions impact rights, livelihoods and reputations, the human must remain firmly in charge. AI may be brilliant in computation, but it must be grounded in law, guided by ethics and supervised by people who understand the weight of public responsibility. Responsible AI is not merely a technical aspiration; it is a moral, legal and administrative necessity. As we integrate AI into governance, our commitment must be to ensure that the technology is safe, trusted and ultimately aligned with the values of our democracy. Only then can AI truly become a force for good—supporting transparent governance, enhancing public trust and strengthening the institutions that serve our citizens. Ultimately, AI must remain grounded in law and guided by ethics, with human judgement firmly in charge—a moral, legal, and administrative necessity for trustworthy governance.



AI for Enhanced Infrastructure Delivery, Transparency & Accountability- A Case of Infrastructure Delivery



Dr. Murali Jagannathan

Assistant Professor, IIT Madras

Introduction-Connecting Vigilance, Infrastructure Delivery, Contract Management and Artificial Intelligence (AI)

The word ‘vigilance’ resonates with good practices such as corruption-free business transactions, transparency and accountability, among other things. All three aspects-business transactions, transparency, and accountability-are crucial and closely interconnected in public procurement processes, and such procurement processes are central to infrastructure delivery. One of the key objectives of the procurement process should be to select the right contractor who can deliver projects within the agreed-upon timelines and budget. However, owing to the nature of construction projects, project deliveries often suffer from time and cost overruns. As of 1st December 2024, the data from the Ministry of Statistics and Programme Implementation (MoSPI), Government of India (GoI), shows that out of 1742 projects being monitored by the Ministry (MoSPI), 882 projects (costing about Rs. 31.03 Lakh Crore) experiencing delays reported a total cost overrun of Rs. 5.01 Lakh Crore in 435 projects (MOSPI, 2024). These numbers indicate the magnitude and impact of time and cost overruns, underscoring the need to implement control measures for improved service delivery.

There have been many past attempts to understand the reasons for time and cost overruns in construction projects (Doloi et al., 2012; Iyer et al., 2008; MOSPI, 2024; PMI&KPMG, 2019) and many delay-causing factors have been identified, many of which are centred around the agreed upon contract document (claims and disputes, adversarial contract conditions, and the drafting of contract documents). Therefore, the solutions to mitigate delays and cost overruns must align with the contractual provisions to ensure better infrastructure deliveries. With the agreed-upon contract being a core element in all financial transactions leading to infrastructure delivery, the role of the vigilance team is indispensable in developing an integral contract management solution for enhanced infrastructure delivery. The points mentioned in *Figure 1* should be the *key focus areas*.

RISK ANALYSIS OF CONTRACT DOCUMENTS

- Thorough Risk Analysis of contract documents before they are released as part of the tender documents
- Benchmarking contract drafting practices with international best practices

AUTOMATED GENERATION OF TENDER DOCUMENTS

- Tender document generation
- Compliance analysis

MINIMISING CLAIMS AND DISPUTES

- Improved correspondence management
- Legal training on contract management for engineers

Figure 1: Key focus areas.

In all the above focus areas, Artificial Intelligence (AI) plays a major role in developing robust contract management solutions.

The Role of AI

The current practice of contract management, which is a manually driven activity, is prone to errors and delays, as evidenced by the sheer size of disputes in the sector (Dubey &

Jagannathan, 2025). With the advancement in AI, many tasks that are otherwise time consuming, repetitive, error-prone or intensive can be automated, pointing to the direct application of AI to solve some of the contract management problems faced by the sector. Among the various applications of AI, tools that aid in human decision-making are one type where AI models assist human decision-makers by providing them with reliable information in a relatively short time, allowing for informed and faster decisions. Natural Language Processing (NLP) is one such branch of AI that involves language- or speech-based tasks, such as analysing textual information and generating text based on specific instructions, as well as translation-related tasks (Balaji et al., 2025). ChatGPT, developed by OpenAI, is a widely used application (Saka et al., 2023). Business operations can greatly benefit from the ‘thinking’ power of NLP models that process language/text, especially when decisions must be made after text mining or sifting through numerous documents and a substantial amount of information. Specifically, in the context of construction contract management, NLP has numerous applications, ranging from simple document analysis to developing chatbots that assist humans in making complex decisions. Some use cases specific to Indian Public Sector Undertakings (PSUs) are discussed in the subsequent sections.

Use Case 1: Risk Analysis of Construction Contract Documents

Construction contracts are legally binding documents that can run to several pages. Typically, a contract document can contain many explicit and implicit conditions that actually transfer risks from one party to the other (Ranjan et al., 2025). When viewed from the perspective of a Public Sector Unit (PSU), such contracts must be carefully drafted to ensure optimal risk transfer, whereby the risks are allocated to the party best equipped to handle them (Subramanyan et al., 2012). In most cases, PSUs tend to use the standard General Conditions of Contract (GCCs), templates that contain many baseline terms and conditions. However, the extent of risks and the capability to handle them can vary from project to project and contractor to contractor. Therefore, to incorporate customisation, Special Conditions of Contract (SCC) are drafted to suit the specific aspects of a given project.

To draft SCCs, a contract drafter should ideally first analyse the standard GCCs, understand the contract terms in the light of project-specific risks, and then modify them through specially drafted SCCs. However, the task of reading a contract and analysing it requires time, knowledge of contract interpretation, as well as knowledge of specific risk areas in the contract based on past experiences. For instance, let us assume that a contract drafter wants to understand the contractual start date for the project. To understand this, a drafter must examine at least three different clauses: clauses related to site availability and its connection with the project start, prerequisites such as the submission of a performance bank guarantee, and the contractual implications of not commencing work on the stipulated start date of the contract. This is one such point for analysis among a hundred different checkpoints that a contract drafter must keep in mind before sharing the document with the bidders for tendering purposes. Not all engineers/officers are equipped to manually perform this task; therefore, contracts can become suboptimal risk transfer mechanisms, leading to delays, disagreements, claims, and disputes in the future (during contract execution). Specifically referring to the Indian PSU environment, it is here that NLP can help in multiple ways, as highlighted below in

Figure 2:

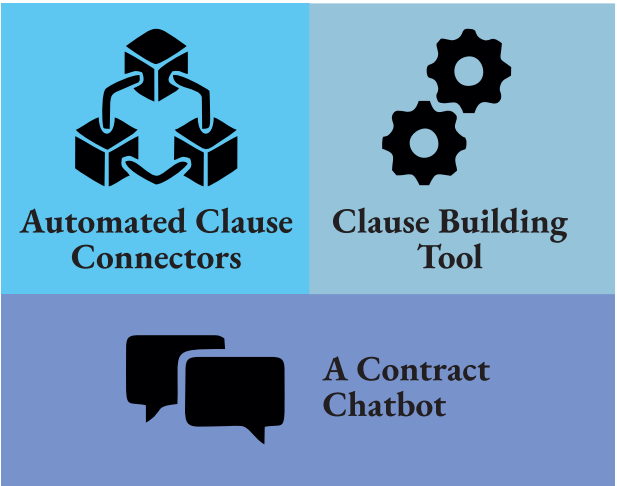


Figure 2: Multiple use cases under “Risk Analysis of Construction Contract Documents

- 1. **Automated Clause Connectors** - A tool that can connect various parts of a contract document and provide all relevant clauses against a set of queries posed by the user. On connecting such clauses, risk analysis can be

performed. Some examples of such rulebased assessments can be found in (Agrawal et al., 2021; Padhy et al., 2021)

- 2. Clause Building Tool for benchmarking with international practices** – This tool can help aid the clause drafter in understanding missing provisions or redundant provisions when compared with a standard reference document. For instance, a clause drafter might be interested to know the extent to which the clause in their document differs from an international practice (like comparing with a FIDIC contract). An example of such similar works, from an international perspective, would be (Lee et al., 2020).
- 3. A Contract Chatbot** – This tool is an extension of the Automated Clause Connector, which can not only answer user-based queries, but also highlight the risk in the clause by linking the interpretation with previous court judgments or even the previous experiences of the organization (information of which is available in the tool’s database).

The use cases and the specific applications listed above are just illustrations, and many such applications can be developed to assist the user in faster and more efficient contract interpretation

Use Case 2: Automated Tender Document Generation

PSU firms typically procure through an elaborate tendering process that typically begins with identifying the necessity for a facility/building/infrastructure, developing a feasibility report, environmental and social impact assessment (depending on the project scope and scale), developing a detailed project report, obtaining administrative approval, preparing estimates, deciding on the procurement mode and then finally developing the tender documents. In fact, all these time-consuming and document-intensive steps preceding the drafting of the tender document contribute to the process leading to the development of a tender document. For instance, if the procurement mode is decided as an item-rate contract, a certain set of conditions and a tender document template should be prepared. If the procurement mode is

going to be design and build, then certain other clauses/ templates may be required. Additionally, depending on the project scope, cost, and timelines, certain policy guidelines should be followed. For instance, depending on the project cost, a certain class of the empanelled bidders may be invited to bid. Similarly, the project cost can determine the inclusion or exclusion of certain provisions, as well as the signing authority, among other things. In simpler words, tenders should be compliant with an organisation’s policy documents and the Government of India’s General Financial Rules (GFR) norms. Any errors in the tender document preparation can lead to disputes later during project execution. Along these lines, once bids are submitted, the bid assessment process is crucial for evaluating bid responsiveness, a process that is manually driven, time-consuming, and prone to errors. In summary, manual checking and the manual preparation of tender documents and bid assessment can be time-consuming and error-prone, and in this regard, the following use cases as shown in *Figure 3* can be useful:

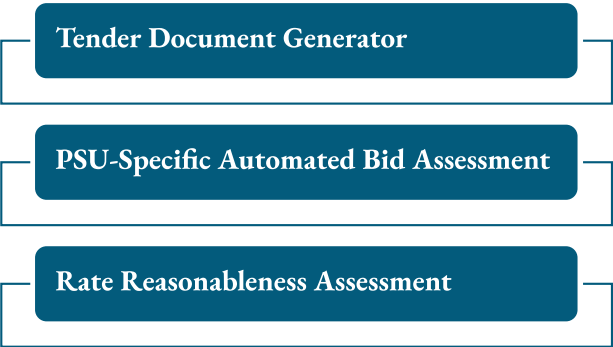


Figure 3: Multiple use cases under Automated Tender Document Generator

- 1. Tender Document Generator:** A tool that is trained with past cases of input (Feasibility reports, DPRs, approvals, policy documents, drawings, etc.) and output documents (tender documents) and can intelligently generate draft tender documents based on the given input documents. An example of research work in a similar area can be seen in (Eken, 2022).
- 2. Automated Bid Assessment:** A tool that can assess a bid’s responsiveness and the bidder’s responsibility, and can quickly provide feedback regarding the assessment of

bids. Some PSUs have already implemented such tools for their procurement processes.

3. Rate Reasonableness Assessment: In many instances, bidders tend to aggressively quote with an aim to secure the contract. It is important to identify such cases and ensure suitable guardrails are in place. This tool is designed to assess the reasonableness of the rates quoted by the bidders by comparing the rate with past executed rates from similar projects.

Again, the use cases mentioned above may or may not be applicable to certain organisations, depending upon their scope of operations. Similarly, many additional use cases can also be conceived depending on the circumstances faced by the organisation.

Use Case 3: Minimising claims and disputes

After the contract award, contract administration is a crucial task that significantly influences the successful completion of a project. At this stage, documentation, responding to contractors' letters, and preparation of contemporaneous records are crucial. However, due to various reasons, if documentation and correspondence fail to meet contractual and legal requirements, it can result in a loss of merit if the matter is disputed at a later date. For instance, in all contracts where time is of the essence, Section 55 of the Indian Contract Act, 1872 requires, among other things, the non-breaching party to intimate the breaching party of the intention to recover compensation. This is important while dealing with extensions of time and liquidated damages. If these actions (intimation to recover losses) are not performed in a timely manner, the PSU may lose its entitlement to compensation. Therefore, if there is a tool that can prompt PSUs to generate certain time-sensitive documents and also create an initial draft based on evolving circumstances (the tool would be trained based on the documentation from past projects). Additionally, it will be beneficial to train engineers to handle legal scenarios through capacity building and training programmes. Furthermore, in the event of disputes escalating to arbitration, it is again a document-intensive process, with documents such as the statement of claims, defence, and rejoinders being prepared by the parties involved in

arbitration. If the documentation is not compliant with the provisions of the contract and the Arbitration and Conciliation Act, 1996, waivers may be invoked, and a PSU may not be able to present its case in arbitration as planned. AI, specifically NLP, can be of assistance. Some of the applications that can help in this are shown in *Figure 4*.



Figure 4: Multiple use cases under Minimising Claims and Disputes

- 1. Automated document generator:** The tool can be equipped to automatically draft contractual letters based on the project scenarios or even incoming contractor letters. These letters can be generated based on certain organisation-specific and legally compliant templates, reducing the massive task of creating the right first draft. The tool can be particularly handy for contract managers who may not have extensive experience (Ranjan et al., 2025).
- 2. Construction court cases summarisation tool:** This tool can summarise complex and landmark court cases without losing the essence of the case. This summary, when validated by experts, can serve as guidance and training to equip PSU engineers for better contract administration (Kshatriya & Jagannathan, 2025).
- 3. Arbitration documentation assistant:** Based on the past cases of arbitration within the organisation, this tool can help in automatically generating the first draft of Statements of Claim/Defence/Rejoinders based on the prevailing circumstances and incoming documents. Also, after receiving the arbitral award, based on the organisation's past experience, the tool can also assist decision-makers in evaluating the repercussions of the decision to either implement or challenge the award.

The above are just a few examples of the various possibilities that can be considered when developing such applications.

Conclusion and the Way Forward

While the above discussions appear useful and interesting, a couple of key points are worth considering. Firstly, since many PSUs handle confidential information, it is essential to consider document confidentiality when developing such tools. This can be achieved by involving domain experts and Data Scientists to ensure data protection. Secondly, the accuracy of all AI models is influenced by the quality and quantity of data available to train the model. If the data

quality and quantity are below the minimum requirements, then the output accuracy can be severely compromised. Therefore, it is first important to understand the available data and its usefulness before fully investing in developing AI models. Next, the developed tool should be free from biases. This will require ‘human-in-the-loop’ and human validation of outcomes. Moreover, it is essential to utilise these models as an aid or an assistant, rather than as an independent decision-maker. The final decision should still be that of a human designated for such purposes. Lastly, it is essential to recognise that every organisation’s needs are unique, and it is crucial to understand the specific requirements before developing any AI-based solutions.



Predictive Cybercrime Analytics



Rajesh Kumar, IPS

Chief Executive Officer

Indian Cyber Crime Coordination Centre (I4C)

Artificial intelligence is rapidly reshaping the landscape of public safety, and in the domain of cybercrime-where criminals are the earliest adopters of cutting-edge technologies-it has become both a challenge and an indispensable ally. At the Indian Cyber Crime Coordination Centre, or I4C, we stand at the intersection of policing, technology, and citizen service delivery. Every day, we witness the scale, sophistication, and speed with which cybercrime evolves. It is against this backdrop that our efforts to build predictive, AI-enabled systems have become central to India’s national cybercrime response.

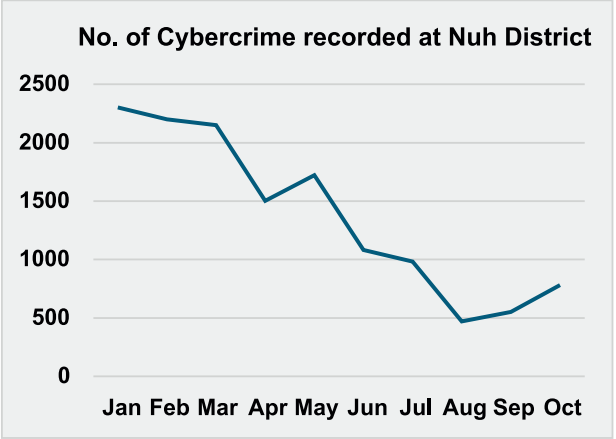
The journey of I4C began with a deceptively simple but deeply consequential challenge: the classification of cybercrime complaints. Through the national cybercrime helpline number 1930, we receive close to one lakh calls every single day. The nature of these crimes spans more than fifty-seven categories-from financial frauds to cybercrimes against women and children, to highly sophisticated cyber-dependent offenses. When classification is incorrect, prioritization suffers. Interventions slow down. Patterns remain unseen. Investigation loses crucial time. Police personnel manning these centers were doing their best, but with such volume, human judgment alone could no longer keep pace.

I4C collaborated through the IndiaAI initiative to develop AI models capable of reading narrative descriptions of

complaints and accurately assigning them to the correct category. These models, trained on vast quantities of real complaint data, now classify with an accuracy upward of ninety-seven percent, enabling us to understand in near real time what types of crime are trending, where they are emerging, and how they are evolving. This capability has fundamentally strengthened our strategic and operational posture.

Yet classification was only the first step. It was soon realized that I4C was up against-highly organized fraud syndicates operating largely from Southeast Asia who were already leveraging AI tools far more aggressively than most public agencies. The scale and precision of their operations demanded a response powered by equally advanced systems. This led to the second major initiative: the complete AI-enabled transformation of the 1930 workflow.

A system was designed where the call-taking agent’s role is simply to converse with the victim. Everything else-voice transcription, translation across Indian languages, detection of key entities, extraction of transaction details, recognition of multimedia evidence sent via WhatsApp or email, validation of bank account numbers, parsing of screenshots, and automatic form-filling-is handled by the AI assistant in real time. The system guides the agent on what to ask next, ensuring that the right investigative information is collected at the right moment. It analyzes the conversation continuously,



listens for fraud indicators, highlights inconsistencies, and fills in all mandatory complaint fields without requiring the agent to manually type or interpret anything. What once took twenty to twenty-five minutes now takes four to five. Errors, which previously caused thirty to thirty-five percent of cases to be escalated incorrectly to banks due to wrong transaction IDs or mismatched details, are dramatically reduced. For banks-many of whom have struggled under the weight of incorrect or incomplete escalations-this system will be transformational.

The next advancement came in partnership with the Reserve Bank Innovation Hub, through a shared ambition to disrupt the financial networks that fuel cybercrime. While banks individually possess visibility only into their own customers and transactions, I4C has a cross-sectoral vantage point that integrates telecom identifiers, device fingerprints, bank account attributes, corporate and GST data, IP addresses, behavioral patterns, and investigative intelligence. This unique ecosystem view enables us to identify mule accounts and fraudulent financial behavior long before any single institution could.



The creation of the Central Suspect Registry allowed us to unify these identifiers into a national repository enriched by AI-generated risk scores. These scores quantify malicious behavior by analyzing both static identifiers-such as PAN, phone numbers, email IDs, and device details-and behavioral indicators such as transaction frequency, velocity of fund movement, anomaly spikes, out-of-pattern activity, or coordinated behavior across networks. When these indicators feed into the Mule Hunter AI deployed across six major banks, its detection capability rises exponentially. What once required months of manual pattern analysis can now be revealed instantly, preventing the circulation of defrauded funds and shielding innocent account holders from being exploited as money mules.

Alongside the financial dimension, cybercrime also exploits vulnerabilities in telecom networks. Every fraudulent call, phishing attempt, or impersonation scam almost always begins with a phone number. To address this, an AI-driven telecom intelligence workflow was created that leverages tower dumps from hotspot districts, which are regions with recurrent cybercriminal activity. These dumps contain millions of records, but AI models sift through them to identify suspicious behavioral patterns. Numbers that make sequential outbound calls, SIM cards that constantly shuffle IMEIs, devices that remain static despite belonging to

A risk score is generated for each device and mobile number. When this score crosses a threshold, district Superintendents of Police are empowered to proactively block the SIM, the WhatsApp account linked to it, the IMEI, and even associated bank accounts before the criminal uses them to victimize citizens.

“mobile” users, numbers with large out-of-state call bursts, SIP lines mimicking legitimate traffic, or SIM boxes generating robotic call patterns-all are instantly flagged.

A striking example came from Nuh district, where major telecom infrastructure disruption led to a rapid drop in cybercrime originating from the region. It demonstrated that when AI, telecom analytics, and law enforcement act together, the results are immediate and dramatic. AI’s role in criminal investigations extends much further. Traditional police workflows involve poring over lengthy FIRs, witness statements, records, transaction trails, call logs, and digital evidence. In complex cybercrime cases, this can become overwhelming.

A system has been introduced that automatically reads an FIR the moment it is uploaded, extracts key events, identifies involved parties, and constructs a chronological narrative that allows the investigating officer to visualize the entire case. The system highlights time-sensitive actions such as obtaining land records, bank statements, digital evidence, or registrar details before tampering becomes possible. It generates interrogation questions tailored to each accused, rather than generic templates, and identifies contradictions, gaps, or inconsistencies in witness testimonies as they are uploaded. It even assembles charge-sheet-ready evidence maps that support prosecutors in building airtight cases.

What would once take days of manual review is now accessible instantly, dramatically improving quality, reducing error, and safeguarding evidentiary integrity. One of the most sensitive areas of our work is the detection and removal of online child sexual abuse material. For this, a sophisticated AI-based tool developed with CDAC Mumbai scans the open web, dark web, and online platforms to detect such content proactively.

It automatically flags, classifies, and reports instances without requiring human exposure to the harmful material. Additionally, it filters evidence files received through international tip lines-particularly from the National Center for Missing & Exploited Children in the United States-to determine whether they contain abusive content requiring urgent law enforcement intervention. Since March 2024, working with platforms under statutory provisions, we have facilitated the blocking of over one lakh URLs hosting illegal content, demonstrating the scale at which proactive monitoring is needed and achievable.

AI agents are also deployed to scrape the internet continuously for unlawful content, including financial scams, phishing infrastructure, impersonation pages, and emerging cybercrime campaigns. These signals are shared with telecom operators, banks, and intermediaries for takedown and coordinated mitigation. The combination of

predictive intelligence, early detection, and rapid action has allowed I4C to move from reactive policing to anticipatory intervention. Through these use cases-real-time complaint classification, AI-enabled call workflows, cross - sectoral mule identification, telecom risk analytics, predictive hotspot monitoring, investigative assistance, internet scraping, and protection of vulnerable communities-I4C is redefining cybercrime fighting in India.

The goal is not to replace human judgment but to empower it. AI augments the intelligence, speed, and reach of law enforcement while ensuring that the scale of cybercrime does not overwhelm the system. The future of cyber policing will belong to nations that embrace this synergy, where human expertise and artificial intelligence work together to secure digital citizens. I4C is committed to ensuring that India leads this future.



Digital Tools in Preventive Vigilance and Good Governance at AIIMS New Delhi



Dr. M. Srinivas

Director, AIIMS New Delhi

In a premier institution such as the All India Institute of Medical Sciences, New Delhi, good governance is not merely an administrative aspiration; it is indispensable to the delivery of timely, equitable and ethical healthcare. AIIMS today manages more than five thousand beds, an annual budget nearing five thousand crore rupees, over fifteen thousand regular staff, and a comparable number of contractual employees. The sheer scale of operations, with more than a hundred decentralized services and verticals, makes the governance landscape both complex and vulnerable. This reality compels constant innovation to ensure transparency and prevent the possibility of malpractice, inefficiency or corruption.

In public discourse, the role of artificial intelligence in healthcare is often associated with clinical advances such as early cancer detection, triaging radiology images, automated laboratory workflows and predictive diagnostics. AIIMS has adopted these technologies widely. Yet, an equally transformative impact of digital tools lies in the administrative sphere, in strengthening internal processes and building an ecosystem of preventive vigilance that operates quietly but powerfully behind the scenes. The first step toward this transformation was recognizing the limitations of legacy systems. While the government's e-Office platform enabled digital file movement, its privilege restrictions, limited searchability and siloed structure hindered comprehensive

oversight. AIIMS' diverse centres—each with its own set of specialized procurement needs—made complete centralization impractical. For example, a pediatric cardiologist requires entirely different equipment specifications than an adult cardiologist, and delays in procurement can affect lives directly. What was needed was not simply central control but an integrated system that preserved necessary autonomy while enforcing transparency, traceability and data-driven decision-making.

This led to the creation of SAHAS—the System for AIIMS Human Resources and Accounting Services—designed and built in-house by our clinicians, IT managers and computer facility teams. It is the first system of its kind in a public hospital of this magnitude. SAHAS unifies three core governance modules: Human Resource Management, Financial Management and the Procurement & Engineering Digital Library. For the first time, HR matters communicate seamlessly with finance, finance communicates with procurement, and procurement links directly with engineering and hospital operations.

The impact of SAHAS on procurement—the area most prone to opacity—has been profound. Every stage of the procurement cycle has been digitized. Requisitions are submitted online, suppliers are selected through automated tools, quotations are compared digitally, approvals are routed electronically, purchase orders are generated instantly and goods receipts and payments are matched automatically. The entire journey of a purchase—from initiation to final payment—is fully transparent and auditable.

A central pillar of this transformation is the Procurement Digital Library, an institutional memory that brings price referencing, rate reasonability, and historical price estimation into the mainstream of decision-making. Departments can now compare prices paid for similar items across the entire institution, while trend analysis reveals fluctuations, demand cycles, and vendor behaviour. The system flags unusual

patterns, such as excessive reliance on particular vendors or prices that deviate sharply from established benchmarks. Internal auditing becomes continuous and proactive rather than episodic. In effect, the PDL functions as the institutional conscience of procurement.

The results have been dramatic. A rationalization exercise on surgical gloves alone yielded savings of fourteen million rupees. Mask prices dropped by almost forty percent. These are not isolated wins but systemic improvements enabled by real-time data and consistent application of procurement principles. The benefits extend beyond cost reduction-AIIMS now has a procurement process that embodies the principles of fairness, competition, transparency, accountability and efficiency outlined in the General Financial Rules. The model also holds promise for replication across public hospitals and medical institutions nationwide.

To extend this framework further, AIIMS is moving toward AI- and machine-learning-driven rate contract management, end-to-end procurement life-cycle analytics, mobile-based monitoring tools for employees, and a fully digital vigilance management system deeply integrated with SAHAS. Unique vendor and procurement codes linked to PAN and GST data are being explored to track transactions across time and departments, while capacity building is being prioritized to ensure that every staff member engaging in procurement understands digital compliance and data governance.

Preventive vigilance is equally important in recruitment, an area where credibility and transparency are critical. AIIMS conducts some of the largest examinations in Indian

healthcare, not only for its own campuses but for more than fifty institutions including all AIIMS, INIs, ESIC, ICMR, VMMC, NITRD, CNCI Kolkata and others. Over seven lakh candidates have participated in these examinations in the past eighteen months, leading to eighteen thousand selections. To protect this process, AIIMS has implemented a tamper-proof, security-enhanced digital recruitment system built on IoT technologies. Facial recognition and Aadhaar-based authentication secure entry into examination centres, while vacancy data, reservation rosters, recruitment rules and preliminary scoring are stored on blockchain, ensuring immutability and trust. Results are declared promptly and the entire process is streamlined through a single-window digital interface.

Beyond procurement and recruitment, a significant transformation has taken place in the patient journey-traditionally one of the most vulnerable areas for exploitation. With thousands navigating the campus daily, intermediaries once offered to “help” with appointments, queue positions, transport within the campus or obtaining priority dates for imaging and tests. These practices were systematically eliminated through digital redesign.

Online appointment systems ensure that no one can charge patients to bypass schedules. Queue management displays in OPDs remove discretionary sequencing by staff.

Real-time public dashboards display emergency waiting times, available beds and OPD loads, making the institution’s functioning transparent to everyone. The Smart Lab, operating 24/7 with robotics and AI, processes more than 1.4



“**AIIMS DISHA, a digital navigation platform, guides patients from building to building, from department to lab, without needing intermediaries.**”

lakh samples and 19 lakh tests every month, enabling same-day reporting for investigations that previously required days and created a fertile ground for agents who misled patients toward private labs. The AIIMS-SBI Smart Card adds a further layer of security by offering cashless, ID-linked transactions that remain safe even if a card is misplaced. This initiative is paving the way for a unified “One AIIMS–One Card” payment system across all AIIMS institutions in the country, linked to broader goals of digital inclusion and faster emergency response.

To close the loop, a QR-code-enabled real-time grievance redressal mechanism empowers patients to report any issue instantly. Complaints automatically reach the appropriate administrative tier, ensuring accountability and timely resolution. This has not only improved service quality but has also transformed staff responsiveness, as every grievance generates an auditable digital trail.

Together, these initiatives-SAHAS, the Procurement Digital Library, blockchain-secured recruitment systems, real-time dashboards, digital navigation tools, smart payments, robotic laboratories and QR-based grievance redressal-represent a cultural shift. Preventive vigilance is no longer an external intervention; it is embedded in the everyday functioning of the institution, enabling AIIMS to operate with integrity, efficiency and compassion.

The transformation demonstrates that a public hospital, even one as vast and complex as AIIMS New Delhi, can uphold the highest standards of governance by relying on data, transparency and well-designed digital ecosystems. As we continue to build on this foundation, the goal is not just to eliminate vulnerabilities but to create a system where trust, accountability and patient welfare become inseparable from the institution’s operational framework.



Transforming Public Service Delivery: Fraud Detection in Government Health Schemes



Dr. Sunil Kumar Barnwal, IAS
CEO, National Health Authority
Ministry of Health and Family Welfare

As India accelerates its transition toward digitally enabled service delivery, ensuring ethical governance across public programmes becomes increasingly critical. The Ayushman Bharat Pradhan Mantri Jan Arogya Yojana (AB PMJAY), implemented by the National Health Authority (NHA), stands as the world's largest government-funded health assurance schemes—today covering more than 15 crore socio-economically vulnerable families. With an annual claim volume that grows each year, the programme demands robust safeguards to prevent fraud, abuse, and leakages.

Globally, up to 15% of healthcare expenditure is estimated to be lost to fraud and abuse. In India, similar trends exist in private health insurance. In contrast, AB PMJAY has recorded a lower fraud and abuse burden—1.2% of total submitted claims under PMJAY (as per FY 2023-24 data). The backbone of this anti-fraud ecosystem is the deliberate and ethical use of Artificial Intelligence (AI), Machine Learning (ML), deep learning, image analytics, OCR, and network analysis to detect, prevent, and deter fraudulent activities across the AB PMJAY value chain.

This article highlights how NHA leverages these emerging technologies to strengthen public trust, accountability, and ethical governance.

The Anti-Fraud Framework

AB PMJAY's anti-fraud ecosystem and its four foundational pillars (Prevention, Detection, Deterrence and Monitoring) are enabled by technology-driven tools that allow NHA and its state partners to respond proactively rather than reactively. This ecosystem developed through a framework, the evolution of which is shown in Figure 1.

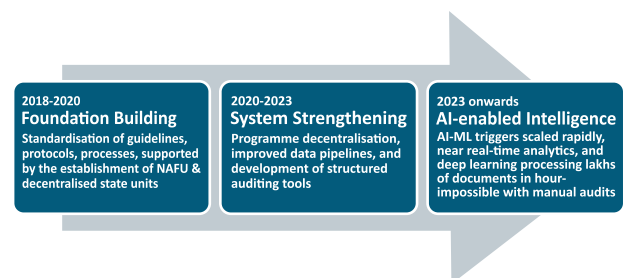


Figure 1: Evolution of anti-fraud framework of AB PMJAY

Through this evolution, NHA moved from weeks or months of delay in detecting anomalies to running all anti-fraud analytics within 24 hours of claim submission.

Fraud Risk in Healthcare and AB PMJAY's Response

Fraud risks in health insurance typically cluster around three actors: beneficiaries, providers, and agents/frontline workers. AB PMJAY has also experienced similar patterns—identity theft, false claims, upcoding, collusion, tampered documents.

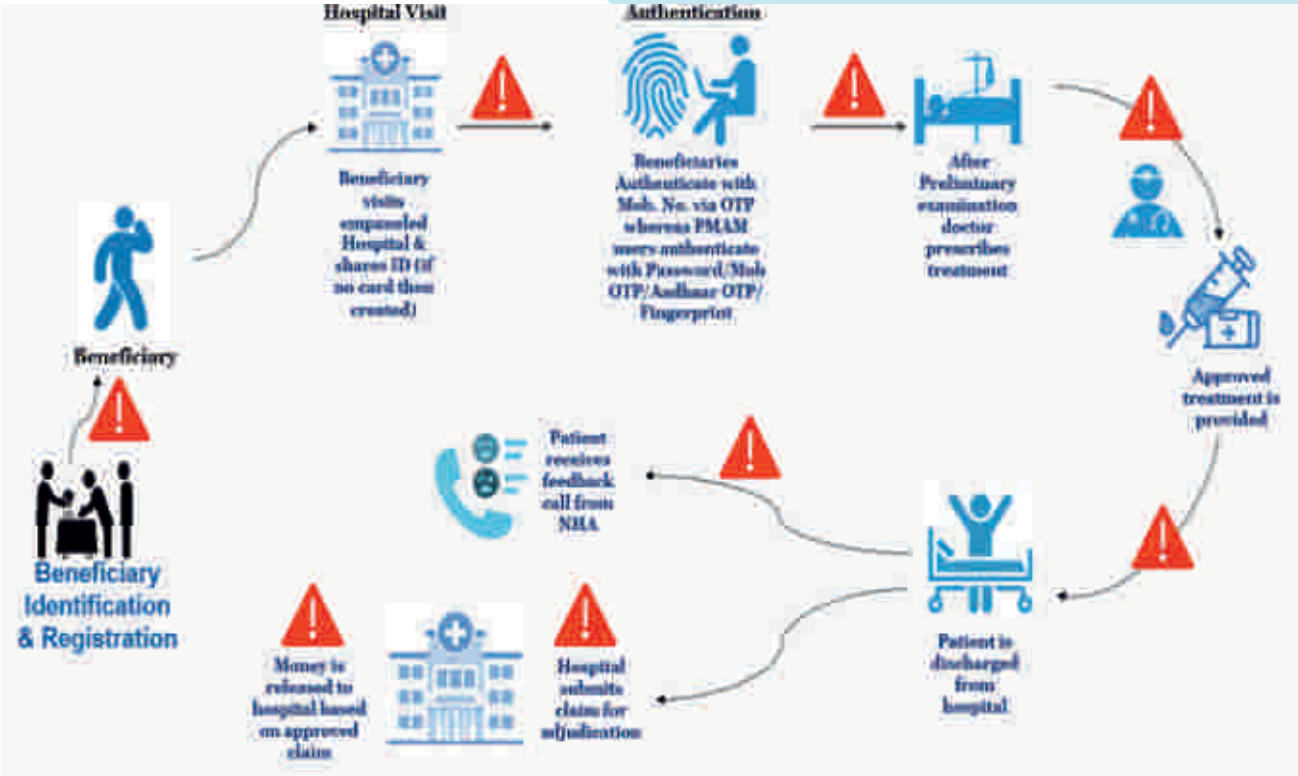


Figure 2: Anti-Fraud/ Abuse checks during Patient Journey in AB PMJAY

AI-Enabled Checks Across the Patient Journey

At every step, as shown in Figure 2, from enrolment to admission, treatment, claim submission, and payment release-AB PMJAY employs a series of automated, intelligent checks. These include bio-authentication verification during beneficiary identification, AI-ML based triggers on card creation and claims processing, and OCR and image analytics for document validation.

This layered approach ensures that fraud is not only detected but often prevented before financial loss occurs.

AI, ML and Advanced Analytics: How AB PMJAY Detects Fraud

At the core of NHA's anti-fraud intelligence are 61 AI-ML and rule-based triggers applied to every claim within 24 hours. These triggers analyse patterns, detect anomalies, and flag cases for further investigation.

Beneficiary Identification: AI for Verification at Scale

To strengthen authenticity of Ayushman Cards, NHA deployed six AI-ML enabled triggers, including:

- Image analytics for face mismatch
- Cross-checking eKYC photos with card creation photos
- Time-based rule analysis for unusual creation hours
- Identification of improbable family patterns (e.g., members from multiple states)

Even auto-approved cards undergo a 24-hour trigger-based review, ensuring continuous scrutiny.

Duplicate & Similar Image Detection

Using deep learning technology, we can identify:

- The same pathology or radiology reports submitted across multiple patients
- Identical pre or post-procedure photos reused to claim benefits

- This technology scans lakhs of images, detecting patterns invisible to the human eye.

Detecting Forged or Tampered Images

AI tools identify manipulated images by examining inconsistencies in:

- Text layers (name, patient ID, timestamps)
- Geotag metadata
- Visual patterns

For instance, identical X-ray images with modified patient details submitted for different individuals are flagged immediately, an example shown in *Figure 3*.

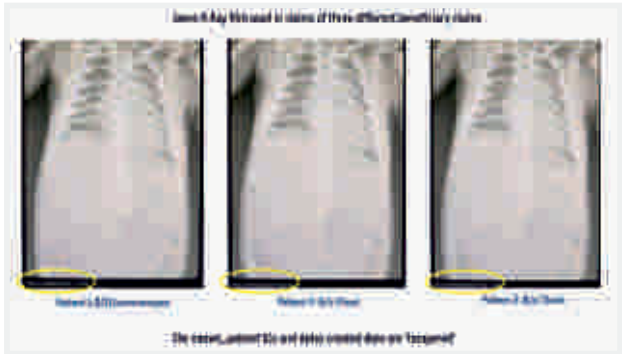


Figure 3: Detection of same X ray films with tampered details submitted for different beneficiaries under AB PMJAY, detected using AI-ML image analytics

OCR and Structured Data Analytics

Optical Character Recognition (OCR) allows the system to:

- Read handwritten lab reports
- Validate clinical parameters with package rules
- Detect statistical anomalies like overbilling
- Identify duplicate invoices using barcode similarity

A common example: a “normal” leucocyte count report submitted for a “severe sepsis” package—flagged within seconds.

Deep Learning for Medical Image Validation

AI models analyse radiological images to verify whether mandatory implants are visible in post-operative films as shown in *Figure 4*.



Figure 4: A patient booked under “Patella fracture K-wire fixation” lacked any visible implant in the X-ray-suggesting a false procedure claim

Network Analysis for Cartel Detection

NHA uses advanced graph-based network algorithms to detect collusive behaviour:

- Shared images across distant hospitals
- Clusters of claims with identical documentation
- Users creating multiple illegal cards used for same-day admissions
- Patterns of coordinated admissions for high-value procedures

These analyses reveal whether fraud is isolated or systemic. An example of such analysis is shown in *Figure 5*.

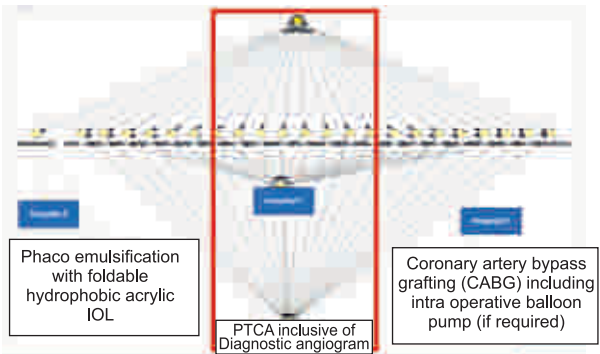


Figure 5: One user (USER3632) created multiple beneficiary cards (20+). All cards were used for admissions on the same day at same hospital. All beneficiaries were registered for the same high value procedure suggesting potential cartel behaviour.

AI for Prevention: Ensuring Ethical Governance

Beyond detection, AI also strengthens systemic governance and prevents misuse.

Integrating National Registries for Accountability

HFR (Health Facility Registry) + HPR (Health Professional Registry) were integrated with AB PMJAY systems to verify:

- Facility authenticity
- Doctor credentials
- Role-based accountability

This prevents unauthorised individuals from manipulating the system.

Role-Based Access and User Sanitisation

NHA implemented a massive clean-up of system users through automated AI checks:

- Reducing users from 7 lakh roles to 5 lakh active users
- Identifying role conflicts (e.g., one individual acting both as verifier and approver)
- Instituting monthly automatic sanitisation

This ensures only authorised and accountable personnel operate within PM-JAY systems.

Hospital Risk Scoring

AI-generated risk scores assist in:

- Prioritising high-risk hospitals for intensive audits
- Monitoring recent, frequency, and financial impact of fraud
- Strengthening deterrence and transparency

Real-Time Dashboards for Governance

Over 50 KPIs integrated into dashboards allow states and NAFU to track:

- Suspicious claim patterns
- Trigger outputs
- Package utilisation
- Geo-spatial anomalies
- Time-based spikes

This creates a culture of continuous monitoring grounded in data ethics and transparency.

Impact of AI and Ethical Tech Adoption

The shift toward AI-enabled analytics has transformed PM-JAY’s anti-fraud performance.

Improved Trigger Efficacy (Figure 6)

- AI-ML contribution to triggers increased from 5% (2022) to 50% (2025)

- Trigger accuracy improved from under 10% to about 50%
- Greater precision in identifying high-risk claims

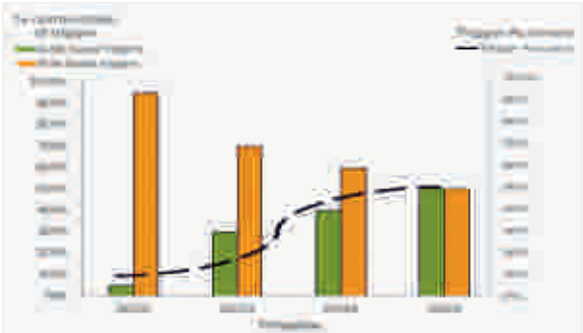


Figure 6: Improved trigger accuracy with increasing use of AI-ML technology

Financial Safeguards

Through AI-enabled detections, PM-JAY has:

- Confirmed fraud worth Rs 630.18 crore
- De-empanelled 1,184 hospitals
- Filed 26 FIRs
- Levied penalties of Rs 231.88 crore

This demonstrates the tangible financial and ethical benefits of AI-led governance.

PM-JAY as a Model for Ethical AI in Public Service Delivery

The PM-JAY experience shows that AI and emerging technologies can drive ethical governance when used with:

- clear guidelines
- human oversight
- transparent, auditable systems
- strong privacy and security protocols

The approach demonstrates that *AI is not a substitute for governance-it is a catalyst that enhances it.*

NHA’s framework emphasises:

- Responsible data use
- Transparency in decision-making
- Fairness in trigger design
- Accountability mechanisms
- Minimal manual discretion
- Prevention of exclusion errors

This makes PM-JAY a leading global example of how emerging technologies can uphold public trust while safeguarding public funds at scale.

Beyond Algorithms: India's Experts Reflect on Trust, Privacy & the Future of AI Systems

Q

What is the authenticity and validity of the results generated by AI? During disciplinary proceedings, when voluminous data is summarized by AI, could crucial facts-potential turning points in the investigation-be compromised?

Answer: Abbishek Singh, Additional Secretary, MeitY, DG, NIC & CEO, India AI Mission

Most AI models-especially those dealing with language-achieve an accuracy of around 85–90% in tasks such as translation. Summarization tools tend to perform even better. However, whenever AI is used in disciplinary proceedings or any official matter, human oversight remains essential. AI outputs cannot be adopted blindly. The technology is known to hallucinate, make errors or present inaccurate facts. Even if AI is used to speed up drafting, such as writing a DO letter, the user must review and verify the content. AI-generated conclusions should always be traced back to the source document to confirm accuracy. While AI-based summarization may be more objective than individual human judgment, every output must still be cross-checked. Human validation will remain indispensable despite the efficiencies AI provides.

Q

What, in your view, will be the next frontier after Generative AI? Will it be quantum AI, autonomous systems, or something else?

Answer: Prof. (Dr.) Anirban Dasgupta, IIT Gandhinagar

Quantum computing is theoretically fascinating, and many of us, including myself, have worked on it. But in practice, only a few quantum algorithms-such as Shor's factoring algorithm and Grover's search algorithm-are genuinely faster than classical ones. Moreover, recent work from UC Berkeley shows that many claims of quantum speed-ups have been matched by improvements in classical algorithms, often inspired by quantum research itself. Therefore, it is difficult to state confidently that quantum computing will be the next major breakthrough in AI usability.

“

In terms of what is emerging right now, AI agents seem to represent the next frontier. Multiple agents interacting to perform complex tasks is an exciting and rapidly evolving space.

We must solve the challenges around AI agents before speculating too far into the future.

”



How does one balance the reliability of AI-generated outcomes with the need to protect privacy?

Answer: *Deepak Goel, Scientist G & Group Coordinator, MeitY*

Privacy is extremely complex because perceptions of privacy differ across individuals. This complexity naturally extends to generative AI outputs as well. Two people may interpret an AI-generated summary or prediction of the same document very differently. The key is understanding the dataset on which an AI model has been trained. If the training data contains sensitive or personal information, then evaluating privacy becomes even more challenging. There is no simple answer; it is a difficult balance to strike.

Additional Response: *Prof. (Dr.) Anirban Dasgupta, IIT Gandhinagar*

There are theoretical approaches to privacy, the most established being *differential privacy*. This ensures that even an adversary cannot determine whether a specific individual's data was included in the dataset. Achieving differential privacy is challenging and often comes at a cost-reduced model performance. Additionally, differential privacy is not always achievable; for some problems, it can be mathematically proven impossible. Thus, many open research questions remain, and privacy continues to be a difficult, evolving challenge.



Can artificial intelligence analyze contract clauses from the perspective of legal tenability?

Answer: *Dr. Murali Jagannathan, Assistant Professor, IIT Madras*

Yes, AI can analyze contract clauses for legal validity. This can be done using rule-based models or knowledge-based systems. For example, ambiguity is a major issue in contract management.

To detect and resolve ambiguity, AI can apply established legal rules such as the order of precedence, blanket clause provisions, invisible signature rules, or contra proferentem principles. A rule-based system can be designed to apply these interpretive principles automatically. Alternatively, knowledge graphs can encode these rules and relationships,

enabling AI to identify whether a specific contractual issue has a defined solution within the encoded knowledge.



As long as rules are clearly defined, AI can assist effectively.



Most cybercrime tools operate through mobile phones. Why are mobile numbers not fully verified with biometrics so that every cybercrime caller can be easily traced?

Answer: *Rajesh Kumar, IPS, CEO, Indian Cyber Crime Coordination Centre (I4C)*

This question is best addressed by the Department of Telecommunications, but I can share that several initiatives are underway. A new pilot system will soon display the caller's name on incoming calls. Facial recognition technologies are being used to verify the identities of individuals purchasing SIM cards through a system called ASTR, leading to the disconnection of large numbers of fraudulent SIMs. Telecom service providers like Airtel and Jio have recently started enabling spam detection, using signals that we at I4C also provide. While the system is not yet perfect, progress is rapid. Biometric Aadhaar authentication is already mandatory for SIM issuance, but fraudsters often use "mules"—individuals paid to provide fingerprints multiple times to obtain SIM cards. We continue to work with telecom operators to tighten these loopholes.



Beyond immediate diagnostic improvements, what is the single greatest ethical or regulatory roadblock preventing scalable AI-driven personalized healthcare from moving from labs into widespread clinical use?

Answer: *Dr. M. Srinivas, Director, AIIMS Delhi*

At present, there are no regulatory roadblocks in India. The only real barrier is the hesitation among people to adopt digital and AI tools confidently. As long as data protection is ensured—and we are already anonymizing datasets extensively—there is no major obstacle. India, with the world's largest patient volumes, is uniquely positioned to train medical AI systems effectively. Large public hospitals provide the perfect environment for AI and ML models to learn from

rich and diverse datasets. This gives India the potential to lead innovations not only domestically but across the entire Global South.



Fraud detection happens after the fraud occurs. What is being done to prevent fraud before it happens?

Answer: Dr. Sunil Kumar Barnwal, IAS, CEO, National Health Authority

This is an excellent question. Most detection currently occurs after treatment, except for beneficiary identification. However, PM-JAY already incorporates an important safeguard: when a patient arrives at a hospital, biometric authentication is performed, followed by a pre-authorization

(pre-auth) review conducted by a remote doctor. The next frontier is to integrate AI directly at the pre-auth stage. For example, if a claim seeks approval for a sepsis package but the total leukocyte count in the diagnostic report is normal, the system will automatically prevent approval. This real-time prevention capability is under development.

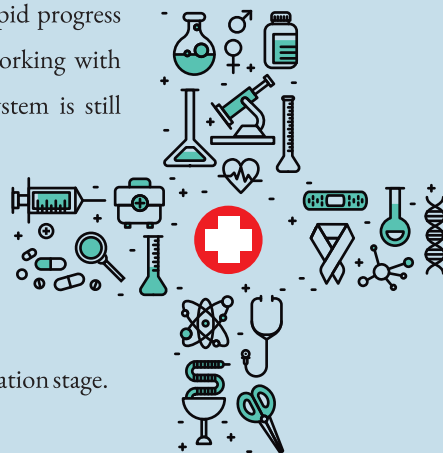
With data from eleven crore treatments, our system is increasingly capable of generating preventive intelligence. We began with rule-based triggers, then brought in vendors to build AI/ML-based triggers, and now we are opening the data to multiple vendors to design advanced AI solutions. The goal is to shift from detection to prevention using predictive intelligence derived from massive historical data.

“Balancing Innovation with Trust and Ethical Safeguards”



he experts agreed that while AI outputs, like summarization, are highly accurate, human oversight is non-negotiable because models can “hallucinate” or present inaccurate facts, especially in official matters like disciplinary proceedings; every AI-generated conclusion must be traced back to its source for verification.

Looking to the future, the next immediate frontier after Generative AI is considered to be AI Agents-multiple agents interacting to perform complex tasks-rather than Quantum AI, whose practical speed advantages over classical computing remain limited to a few specific algorithms. Balancing AI's reliability with privacy is extremely complex, as perceptions differ and models are trained on varied data; while differential privacy offers a theoretical solution, it often reduces model performance and isn't always mathematically achievable. In the legal sector, AI can effectively analyze contract clauses for ambiguity and legal tenability using rule-based systems or knowledge graphs, provided the governing legal principles are clearly defined. Addressing cybercrime, government agencies are making rapid progress using systems like ASTRA (facial recognition for SIM verification) and working with telecom operators to close loopholes exploited by “mules,” though the system is still evolving. Finally, in healthcare, the primary roadblock to scalable AI adoption is not regulatory, but the public's hesitation and lack of confident adoption; however, India is uniquely positioned to train world-class medical AI systems using its large, diverse patient datasets, with the goal of shifting the focus from post-event fraud detection to real-time fraud prevention through predictive intelligence integrated at the pre-authorization stage.



Session-II: Panel Discussion with Experts from Industries



Sh. Nitin Chugh is the Deputy Managing Director and Head of Digital Banking & Transformation at State Bank of India. He has previously worked as the MD and CEO of Ujjivan Small Finance Bank. He holds degree in Engineering & Management.

Sh. Nitin Mishra is the Executive Director, Tech., National Payments Corporation of India. He has previously worked as a CTO with ONDC & GSTN and has extensive experience in design, architecture, development and operating internet scale applications and infinitely scalable technology stack for diverse use cases. He holds a PGDM degree from IIM Kozhikode.



Sh. Mohit Kapoor is currently the Group Chief Technology Officer at the Mahindra Group. He has previously worked with DBS Bank as the head of technology optimization and head of Asia Hub. He is also a member of AI for India-2030 Council at the World Economic Forum.

Sh. Manish Malik is currently the Executive Director (Information Systems) at IOCL. He has been instrumental in the development of various Cloud based Business transformation projects, deploying largest SD-WAN Network for Retail business, introducing Paperless culture by setting up Enterprise Content Management System, digitalization of Supply Chain Systems, GenAI based tender evaluation and multiple IT infrastructure modernization projects.



Sh. Rahul Himalian is an officer of the India Railway Traffic Service and is currently the Director (Tourism & Marketing) at IRCTC Ltd. He is also the chairman of IRCTC Payments Ltd. He holds a Bachelor's degree in Electrical Engineering and a Master's degree from IIT Delhi.

Ms. Geeta Gurnani is Field CTO for IBM India & South Asia, where she leads cross-functional technical pre-sales across IBM's core technology portfolio-Data, Automation, Hybrid Cloud and Infrastructure Platforms for Clients and Ecosystems Partners.



Sh. Dhrubabrata Ghosh Dastidar is the Managing Director at Protiviti India, where he leads the firm's Data analytics and Digital Transformation practice. With a career spanning across India and the Middle East, he has spearheaded large-scale transformation programs across sectors including financial services, technology, media and e-commerce. He is also an alumnus of IIT Kharagpur and ISB.

Leveraging AI and Emerging Technologies for Operational Efficiency & Preventive Vigilance

Session II brought together industry leaders and technology experts for a panel discussion on how AI and emerging technologies are transforming operational efficiency and preventive vigilance. The session highlighted how organizations have been shifting from manual oversight to intelligent, data-driven systems that enable real-time monitoring, automated compliance checks, and predictive risk identification. Panelists shared sector-specific use cases—from financial services and energy to manufacturing and services—illustrating how AI has helped embed vigilance into everyday operations, reduce leakages, and enhance transparency. The discussion also underscored the importance of ethical AI, responsible innovation, and cross-sector learnings in building trustworthy and future-ready vigilance frameworks.



Q SBI has been at the forefront of digital banking with platforms like YONO. With GenAI and Agentic AI now being adopted widely across industries, what are some prominent AI use cases SBI or the financial sector is currently working on?

| Nitin Chugh, Deputy MD, SBI

At SBI, we have deployed more than 150 AI-driven use cases across operational efficiency, fraud and risk management,

revenue leakage prevention, customer service, and business intelligence. Although we are actively exploring generative AI, we have intentionally restricted customer-facing deployments until our internal validations are complete. Internally, we have already operationalized seven GenAI models, one of which is available to all employees.

Traditional AI continues to be extremely valuable. Our ensemble model for identifying high-risk branches, which we have patented, brings together fraud data, compliance

patterns, regulatory observations and people-related parameters. This allows our controllers and auditors to walk into branches with a clear view of potential issues rather than starting from scratch. Our revenue-leakage and SME pricing models have also delivered strong results. AI is deeply integrated in our operations.

Q AI relies heavily on high-quality data. How have you managed challenges arising from legacy datasets?

| Nitin Chugh, Deputy MD, SBI

We addressed our legacy data challenges through Project Ganga, an eight-year data cleansing and standardization exercise. We no longer face the typical problems associated with missing or inconsistent data.

Most of our models rely on transaction data rather than PII, which helps with compliance as well as quality. We typically train models on the most recent two years of data, and this has been more than adequate for building reliable & effective systems.

Q NPCI handles some of the largest digital transaction volumes in the world. What AI use cases are emerging in payments and other related sectors?

| Nitin Mishra, ED (Tech), NPCI

NPCI processes nearly 650 billion transactions every month, so our digital-first architecture has naturally prepared us for AI adoption. While we remain cautious because of the sensitive and regulated nature of financial data, we are now building several strategic AI models.

Recently, we introduced Small Language Models (SLMs) for dispute resolution; we showcased these at the Japan FinTech Festival. These models perform exceptionally well because UPI and other NPCI platforms generate clean, structured trace data. We are using AI for cybersecurity threat detection, fraud pattern prediction, code analysis, user experience improvements and predictive outage detection.

Foundational LLMs will continue evolving, but in my view, SLMs will dominate sector-specific deployments going forward—they are efficient, scalable and purpose-built.



Do you believe SLMs will see wider adoption than large foundational models?

| Nitin Mishra, ED (Tech), NPCI

Yes. Foundational models will evolve incrementally, but the real innovation will come from lightweight, domain-specific SLMs. These models are easier to deploy, consume fewer resources, and perform exceptionally well for well-defined industrial tasks. They will increasingly replace repetitive human work and power new conversational and decision-support systems.



These days companies operate in diverse sectors. How are you using AI to enhance efficiency across such a wide range of business verticals?

| Mohit Kapoor, Group CTO, Mahindra Group

We use AI extensively across all our business lines. For instance, we handle over 200,000 monthly emails, and GenAI agents now categorize, summarize and route many of them directly into our CRM systems. On the shop floor, we've introduced WHY-GPT, a diagnostic GenAI tool that helps supervisors interpret faults using image recognition and historical repair data.

Dealership engineers now use AI tools that convert customer complaints—such as “khat-khat sound in my Thar”—into specific diagnostic insights. We also launched Vehicle GPT, which replaces the entire car manual with an AI assistant.

In agriculture, we combine weather, soil, satellite, subsidy and mandi data to improve decision-making for farmers. In our financial services vertical, AI models help us assess intent-to-pay and determine optimal loan rescheduling windows. Across the Group, we track time saved for employees and customers as a core KPI.



Are you also working on physical AI, such as humanoid robots, for industrial operations?

| Mohit Kapoor, Group CTO, Mahindra Group

Not humanoids in the social-media sense, but we use extensive physical AI. Our new Chakan plant uses over 900 robots, not

only for assembly but also for performing quality checks. These robots work collaboratively—one installs, another seals, a third inspects—reducing the need for manual intervention and improving consistency and safety.



India's linguistic diversity is a major challenge. How are you building vernacular capabilities into your AI models?

| Mobit Kapoor, Group CTO, Mahindra Group

This is a complex challenge. We are addressing it through Project INDUS, a major joint initiative with Tech Mahindra to support 15 Hindi dialects and other regional languages. We scanned physical books from national and state libraries to train these models. Project INDUS aligns with the government's goal of building sovereign Indian AI models that can function inclusively across linguistic diversity.



IOCL manages a massive energy network. Beyond predictive maintenance, what emerging AI use cases are you seeing across refineries, pipelines and retail?

| Manish Malik, ED (CIS), IOCL

Predictive maintenance has transformed our operations. Sensors monitoring temperature, vibration, pressure and acoustics have significantly improved our ability to anticipate failures, especially for mission-critical components like FCCU blowers. With annual R&M spending of ₹7,000–8,000 crore, even a 10% improvement yields enormous savings.

We generate large volumes of data across LPG, retail, CRM and ERP systems, but much of it is still scattered. Our first priority is building a unified data lake to serve as the foundation for future AI deployments.

In pipelines, we are using distributed acoustic sensing through optical fibers. Our ML models can distinguish between digging, hammering, vehicular movement and pumping noise, and we overlay this with socio-economic, geographic and crime data to identify high-risk areas. This allows us to respond instantly and prevent pilferage or leakage.



How does data consolidation shape IOCL's broader AI roadmap?

| Manish Malik, ED (CIS), IOCL

Without good data infrastructure, AI cannot scale. Consolidation is our immediate priority. Once the data lake is robust, we will expand our use cases across marketing, logistics, refineries and pipelines in a much more coordinated manner.



IRCTC runs one of the world's largest ticketing systems. How has your organization moved from manual to digital to AI-powered operations?

| Rabul Himalian, Director (T&C), IRCTC

Our journey has been remarkable. From long queues and paper tickets, we have moved to a stage where 87% of all railway tickets are booked online—that's about 14–15 lakh tickets daily. During peak Tatkal windows, we handle up to 37,410 tickets per minute.

AI/ML now plays a crucial role in three areas:

- 1. 139 Integrated Railway Helpline:** We manage 2.5 lakh daily calls and 1 lakh SMS across 13 languages. With Azure OpenAI, Meta Llama, TTS, ASR and NLP, our IVR now resolves 89% of queries, reducing abandoned calls from 30–40% to just 1%.
- 2. BOT Detection & CDN Optimization:** AI-based CDN offloading has reduced PRS/NGET load by 90% in hits and 67% in volume. BOT mitigation now reaches 30–40%, and we have deactivated 2.7 crore fake IDs.
- 3. AskDisha Chatbot:** Handling 1.33 lakh daily interactions, AskDisha now books, cancels and manages tickets. It books 2,500 tickets daily—equivalent to 8.7 lakh annual bookings.

We also apply AI for e-catering (sentiment analysis, spam detection, personalized recommendations), where daily orders have risen to 1.15 lakh. Similar AI tools support Rail Neer plant operations and Heli Yatra bookings, where device fingerprinting helps detect fraud.

Our philosophy remains simple: “*I Really Care Towards Customer (IRCTC)*,” and AI is helping us deliver exactly that.



You bring a very different perspective as an OEM. You’ve seen AI use cases across multiple industries. Could you share key AI/ML applications you’re seeing in India and globally?

| Geeta Gurnani, CTO, IBM

Thank you for having me—it’s been enlightening listening to the diverse experiences shared today. In many ways, it felt like attending a masterclass in industry applications—from railways to energy, BFSI to automotive. Before talking about use cases, I want to highlight a simple yet powerful thought that our Chairman often reiterates: *before adopting technology, rethink the process itself.*

We encourage clients to follow three steps: Eliminate, Simplify & Automate.

- 1. Eliminate** – Ask whether the process even needs to exist.
- 2. Simplify** – If it must exist, reduce a 20-step workflow into four.
- 3. Automate** – Only then bring in AI/ML, GenAI, SLMs, or frontier models.

When viewed through this lens, AI use cases fall broadly into three categories:

- 1. Customer Experience:** This is the biggest area of transformation. Much like the examples shared today—call center automation, email assistance, multilingual bots, personalized messaging—organizations across sectors are using GenAI to make customer interactions more intuitive and responsive. Any department dealing directly with customers should start here.
- 2. Horizontal Functions (HR, Finance, Procurement, IT):** These are common across all large organizations. At IBM, HR has seen the highest productivity gains through AI-enabled query resolution, talent operations, and employee support. Procurement and finance are witnessing major shifts too—GenAI is now used to scan

large contracts, identify inconsistencies in payment terms, automate invoice reconciliation, and detect anomalies. After ERP systems, this is the next big leap in operational efficiency.

- 3. Core Industry-Specific Processes:** This is where AI’s real power shines. For example, one of India’s largest four-wheeler manufacturers had nearly 2,000 engineers manually verifying technical drawings for every spare part—from seat belts to seats. Using GenAI vision models, we reduced a 5-7 day process to just five minutes, while enhancing accuracy and regulatory compliance. Similarly, health insurance faces challenges with handwritten prescriptions, lab reports, and medical claims. AI is helping underwrite policies, detect fraud, and automate verification in areas that were earlier considered “too complex” for automation.

Across industries, these three pillars—customer experience, horizontal functions, and core business processes—are where AI is delivering



A follow-up: beyond Generative AI and Agentic AI, do you see any major technological breakthroughs emerging?

| Geeta Gurnani, CTO, IBM

Predicting the future of technology is always tricky, but one development is very clear—the rise of quantum computing. We are entering what we call the *quantum utility era*. India will get its first IBM quantum system next year in Amaravati through a collaboration between IBM, the Andhra Pradesh Government, and TCS. Many assume quantum is about “speed,” but that’s a misconception. Quantum is about handling thousands of variables simultaneously—something classical systems cannot do efficiently. This is crucial for areas like real-time fraud detection, high-dimensional optimization, and molecular simulation. I foresee a future where AI/ML converges with quantum to solve some of the hardest computational challenges. Regarding GPUs—no, they’re not going away. We will instead move toward hybrid computing: traditional compute, GPUs, LPUs (Language Processing Units), and quantum systems, each chosen based on the business problem. And as Nitin (NPCI) said, this also aligns with the rise of Small Language Models—purpose-built,

efficient, domain-specific models rather than massive general-purpose LLMs.



You deal with huge amounts of sensitive customer data. What is SBI doing in the area of ethical and responsible AI?

| Nitin Chugh, Deputy MD, SBI

At SBI, we have had a Responsible and Ethical AI Framework for five years now-approved by our Board and aligned with global standards including Singapore MAS's FETA principles (Fairness, Ethics, Transparency, Accountability).

Our practices include:

- Model developers and model validators are always separate teams.
- Models undergo calibration and recalibration cycles.
- Deployment happens only if accuracy thresholds are met.
- Outlier behaviour or unexpected results trigger model rollback.

With GenAI, we realized process-based governance alone is not enough. So we automated guardrails in our internal GenAI system *AskSBI*, which is trained on over 14,000 internal circulars.

Employees can ask natural-language questions-but we had to put restrictions. Earlier, people asked, "Why isn't my boss approving leave?" or "When will I be promoted?". We built filters to prevent such queries. As we move toward real-time fraud detection, DPIIP-based risk scoring, and agentic workflows, we need automated guardrail systems that continuously monitor and flag deviations. Ethical AI is not just policy-it is infrastructure, automation, risk management and explainability.



Bias is often a concern in AI models. Do you encounter bias in the models you develop, and how do you mitigate it?

| Nitin Mishra, ED Tech, NPCI

Bias, overfitting, and underfitting are inherent concerns in any statistical model. That's why we adopt multi-level testing:

multilayered hypothesis validation, iterative tuning, rigorous model evaluation using precision, recall, F1 scores, etc. UPI data is high-quality and well-annotated-it's essentially a request and response transaction. Bias risks emerge only when we combine payments data with external datasets (e.g., e-commerce or healthcare). Since we haven't entered that space yet, bias remains manageable. Still, we remain extremely cautious. Questions framed differently can lead to different interpretations by a model, and we ensure our testing processes prevent unintended skew.



Does the nature of your payments data inherently introduce bias?

| Nitin Mishra, ED Tech, NPCI

Not inherently. Payments data is structured and clean. The challenge only arises when it interacts with third-party datasets, where context may create unintended correlations. We haven't gone down that path yet. But as India moves toward sovereign AI models, domain-specific models, and SLMs, we will need rigorous validation to ensure that bias doesn't creep in. We will only deploy models after exhaustive testing.



How do you balance innovation with ethical AI at Mahindra Group?

| Mohit Kapoor, Group CTO, Mahindra Group

Let me start with something interesting related to bias. Most cars in India are driven by men, but buying decisions are often made by women and children. If your training data reflects only the drivers, you'll design a car that appeals to men and alienates women. This is not an AI problem-it's a business insight problem.

So we use an ERS framework-Ethics, Responsibility, Security. We also identify where *not* to use AI.

For example: roadside assistance. If someone is stranded on a highway, they don't want to talk to a bot. We use AI to prioritize such calls so they are instantly routed to a human, even interrupting agents handling routine queries. AI, in this case, helps us *make humans more human*.

Data privacy is also non-negotiable. We host GenAI models in a local, secure container with a “model garden” trained on internal, verified data. Most employee questions-like travel norms, reimbursement, procurement-are best answered from internal documents anyway.

Finally, explainability is critical. If an AI suggests a travel entitlement, it must provide the original policy PDF as citation. This keeps trust high and prevents blind acceptance.

Q How does IOCL ensure ethical, responsible AI across such a large and complex infrastructure?

| Manish Malik, ED (CIS), IOCL

Ethical AI essentially means that AI must respect human values. In traditional software, 2+2 was always 4. In AI, 2+2 can become 3 or 5 if the model is trained incorrectly. That’s where risks begin.

We use AI for contractor evaluations, vendor assessments and predictive maintenance-with 80–90% accuracy. But when the model recommends something, I am accountable for explaining it.

So we must ensure: fair and unbiased training data, equal opportunity (e.g., no gender bias in application shortlisting), strong privacy frameworks aligned with the DPDP Act, and explainable outputs.

We are building an enterprise-wide ethical AI framework with a steering committee (legal, safety, AI experts), a central AI governance layer and checks for anonymization, consent, fairness and role-based access. As we scale our AI/ML platform and data lake, every model will pass through ethical validation before deployment.

Q Any closing thoughts on ethics, governance or the future of AI?

| Rabul Himalian, Director (T&C), IRCTC

AI always brings poetry to my mind. So here’s a sher: “Main safar akela chalta tha... log judte gaye, aur karwaan bantā gayā”. That is exactly what AI and ML do-data keeps coming, patterns keep forming, intelligence keeps building. On ethical governance, there are three layers.

1. Intrinsic Controls: We must not wait for external entities. IRCTC’s systems-firewalls, web application security, data centers and 24×7 monitoring-are intrinsic safeguards. With the scale we handle, responsibility comes from within.

2. Ecosystem Collaboration: We work closely with NCIIPC, CERT-In, CDAC and other national cybersecurity bodies. Our systems feed threat, attack and anomaly data to them, and we act on their alerts. Cybersecurity today is a shared responsibility.

3. Resilience and Innovation: Many challenges will never be documented in a manual. For example, just before Diwali, we had to deliver 11 crore emails to users while our systems were built for 1 crore. We achieved an 11x productivity jump within 36 hours.

Innovation and efficiency together define success-AI is a partner, not a replacement.

Q Finally, from a governance standpoint, how do you address accountability, bias and explainability as a platform provider?

| Geeta Gurnani, CTO, IBM

Governance is critical because decision-making is increasingly being delegated to AI systems. Even if humans remain in the loop for final approvals, AI-driven “straight-through” processes require accountability and auditability. We look at governance through three lenses:

1. Model Lifecycle Governance: Every department wants to innovate. That’s good-but all models must fall under a common governance umbrella with rules for onboarding models, updating them, retiring them and choosing between large or small models.

2. Runtime Risk Management: This includes bias detection, fairness checks, hallucination thresholds and context-specific risk rules. This must be automated because manual oversight cannot scale.

3. Compliance and Regulatory Alignment: Global frameworks (like UAE and Singapore AI Acts) and sectoral guidelines (RBI, IRDAI, etc.) require continuous

compliance. Manual compliance tracking is impossible at scale.

This is why IBM built Watsonx governance, a platform that automates governance, risk and compliance across AI models.

What we tell clients is simple: Don't scale now and fix governance later. Start governance first if you want responsible, trustworthy AI.

And across everything-governance is now inseparable from AI security, because model training, data sourcing, and external dependencies all bring new risks.

Q You have seen global rollouts of AI. What does preventive or agentic vigilance mean to you, and can you share a few examples?

| Geeta Gurnani, CTO, IBM

I must confess-I was intrigued by this question. When it was shared with me, the first thing I did was ask ChatGPT what "preventive vigilance" means. It had no answer! It only asked me for context. So I dug deeper and looked at the root of the word "vigilance," which comes from a Latin term meaning to be *aware or watchful*. But even then, defining *preventive vigilance* wasn't straightforward. For me, preventive or agentic vigilance depends heavily on the industry and the context. For example, when Rahul and I were chatting earlier about railways, he explained the challenges in reservation fraud-where demand far exceeds supply. Their next-generation PRS will need to detect suspicious booking patterns, identify unfair practices, and ensure that genuine passengers get a fair chance. That, to me, is preventive vigilance-*continuous monitoring to ensure decisions are taken in the right spirit*. In BFSI, preventive vigilance is taking shape through AI-driven case management. Complaints, grievances and escalations often carry human bias-fatigue, subjectivity, or past experiences influence decisions. AI systems can remove this bias by ensuring fairness, consistency, and speed. Human teams then focus on judgment, while AI handles the repetitive comparison tasks-just as I mentioned earlier with automobile design verification. A design engineer's role is not to manually compare specifications for days-it is to think creatively and

validate correctness. So to summarize: any area needing continuous, fatigue-free, bias-free monitoring is a strong candidate for preventive vigilance. And I'm sure next time someone asks ChatGPT for this definition, it will have a better answer!



As someone dealing with massive data flows at SBI, how do you view AI/ML in preventive vigilance and fraud detection?

| Nitin Chugh, Deputy MD, SBI

At its core, AI is simply data-and preventive vigilance relies on consuming large volumes of data intelligently. Traditionally, vigilance reacts to signals. But preventive vigilance requires us to *actively generate, interpret and act on signals*.

That's why I see the future of vigilance moving toward something similar to a Security Operations Center, but for governance-a Vigilance Operations Center that continuously ingests data from branches, functions, vendors, and transactions, analyses it for emerging patterns, identifies hotspots, and alerts teams in near-real time. This is similar to how real-time fraud detection works in payments today. Once people know such a system exists, it becomes a strong deterrent in itself. Another important aspect is convergence. Audit, risk, compliance, and vigilance all draw from overlapping data sources. With modern GRC platforms-including the ones IBM demonstrated to us-the intelligence generated for risk management becomes extremely valuable for vigilance, and vice versa. These boundaries are blurring, and AI will accelerate that convergence.



You process massive volumes of high-velocity transactions. How do you see vigilance integrating with AI/ML?

| Nitin Mishra, ED Tech, NPCI

For us, the challenge is not volume-it is velocity. With UPI authorising transactions within a fraction of a second, preventive vigilance must operate in real time. The key lies in three pillars:

1. **Signal extraction in near-real-time:** We must identify meaningful patterns amidst billions of datapoints.

2. **Strong feedback loops:** Today, many fraud investigations end without feeding outcomes back into the model. For preventive vigilance, the system must learn from every closure.
3. **System self-awareness:** Our dispute resolution systems already flag anomalies in seconds, but fraudsters constantly evolve—from social engineering to remote access attacks to phone mirroring. Preventive vigilance requires systems that both detect and self-improve. Real-time risk scoring is already happening at our switch layer for every UPI payment. The next step is embedding continuous learning into these systems.



Do you see vigilance becoming agentic or predictive rather than largely rule-based?

| Mohit Kapoor, Group CTO, Mahindra Group

Absolutely. To me, two ideas stand out.

1. **Moving from Sampling to 100% Coverage:** Thanks to cloud and AI, we no longer need sample-based audits. Every transaction, every reimbursement, every purchase order can be checked—*all the time*. That’s already what we do. Our Agentic AI compares POs, invoices, GRNs, specifications, approval chains and flags exceptions instantly. Payments move faster, leakages reduce, and vigilance becomes continuous, not episodic.
2. **Catching anomalies during the act, not after:** Agentic vigilance means embedding rules directly into systems so fraud is prevented before someone files a complaint. This mirrors how all our vehicles and tractors are connected today. If a quality issue arises in Guwahati, we know instantly—we no longer wait for six-month audit cycles. So yes, vigilance is shifting from post-facto detection to real-time deterrence.



How is preventive vigilance through AI evolving in IOCL?

| Manish Malik, ED, CIS, IOCL

One incident always stays with me. A young girl at a retail outlet used her phone and a plastic bottle while refueling—

both violations of safety SOPs. A fire broke out, and she tragically lost her life. A preventive AI system using video analytics could have saved her. Our CCTVs already capture movements; an AI model can easily detect unsafe behaviour—phones, bottles, unauthorised containers—and instantly shut the nozzle. That is preventive vigilance in action: *acting before the hazard unfolds*.

We’re applying similar models across operations like detecting domestic-to-commercial LPG diversion through booking pattern analysis, identifying fraudulent OTP clusters, using GenAI for tender evaluation to avoid interpretation bias, verifying contract terms using LLMs to flag deviations from SOPs. These systems protect revenue, uphold transparency and save time—allowing teams to focus on complex judgments. If AI prevents even one major fraud, the value far exceeds the cost of deployment.



What do you see as the future of AI/ML in IRCTC and beyond?

| Rahul Himalian, Director (T&C), IRCTC

Agentic AI is powerful. Imagine a recruitment process handled by three agents: one shortlists candidates, another conducts interviews, a third negotiates salary—all reporting to a super-agent. But the real question remains—*will I choose the candidate the agent approves?* That’s where human judgment remains irreplaceable. In IRCTC, the main target for vigilance is “touts”—people blocking genuine users. We’ve blocked 2.5 crore IDs, revalidated 40 lakh, and shut down 12,000 fraudulent email domains. That is our preventive vigilance landscape. AI will help, but so will humane systems. I often use the “headlight approach”—if both drivers keep high beams on, an accident is inevitable. If one dims, harmony prevails. Similarly, organizations must balance AI with empathy, compassion and fairness.



What about the environmental impact of the exponential growth of AI?

| Geeta Gurnani, CTO, IBM

I’d like to reframe the question. When computers arrived, people feared job loss. Instead, new skills emerged. AI will be

similar. People won't lose jobs because of AI—they'll lose jobs if they *don't know* AI. We need massive reskilling. And yes, AI models require compute and energy, but efficiency is improving.

India is planning a major "Impact AI" summit early next year to examine exactly these questions. The real question is: how do we prepare people and processes for this shift? Skilling, governance and responsible deployment are the answers.



How will AI impact the petroleum and energy sector?

| Manish Malik, ED, CIS, IOCL

AI will not eliminate jobs—it will change them. Just as computers did 30 years ago. For IOCL, AI will be essential across EV infrastructure, hydrogen, battery swapping, non-fuel retail, customer segmentation, predictive maintenance, and personalization and engagement. Our profitability lies increasingly in non-fuel offerings, and AI helps us understand customer behaviour better.

We also plan to leverage AI in data centers, land monetization, and new business models. Much like ERP transformed us years ago, AI will become a platform generating hundreds of new use cases. Our younger workforce already thinks digitally—they want analytics, automation, and intelligent systems. AI is a natural evolution.



Beyond AI, what emerging technologies will IRCTC leverage?

| Rahul Himalian, Director (T&C), IRCTC

Our biggest upcoming initiative is the Unified IRCTC Portal. Today we operate multiple portals—ticketing, tourism, HeliYatra, Maharajas' Express, FTR, and more. We're working on a unified, AI-enabled, experience-driven platform that offers personalised tourism suggestions, interactive videos, testimonial-driven recommendations, seamless navigation across services, and deep customer engagement. With 11 crore registered users and 14 lakh daily customers, this will take IRCTC far beyond its current digital footprint.



How can India build a public-private-academia framework to compete globally in AI innovation?

| Geeta Gurnani, CTO, IBM

This is already beginning. India announced a call for "Indian Foundation Models," and startups like Sarvam AI and BharatGPT emerged as frontrunners. IBM partnered with BharatGPT to address a key challenge: India lacks high-quality textual datasets because much of our language diversity exists in voice form.

Most global LLMs are trained on data from only six countries, leading to significant bias. We brought our research capabilities to help train India-specific models efficiently and affordably using Indian datasets. The IndiaAI Mission has played a transformative role by uniting startups, IITs, global companies and government. The first problem being tackled? **Grievances**—a domain every Indian understands. It mirrors the vigilance context where inputs are unstructured and emotionally loaded. This collaborative ecosystem is exactly how India will build competitive, responsible AI.



What is the future of AI across organization, and do you see it as a threat to jobs?

| Mohit Kapoor, Group CTO, Mahindra Group

In autos, AI's biggest promise is safety—predicting accidents, coordinating traffic signals, improving braking systems, and saving lives. The ecosystem—not just the vehicle—must improve together. Our vision is global-quality features at Indian affordability. On jobs—yes, AI poses a threat if we do not upskill. Automation transformed factories a decade ago; AI will transform digital work now. But we invest heavily in workforce development. Another big focus area is rural and semi-urban India. People have smartphones, but digital public goods are still not fully accessible. We are committed to making livelihoods stronger—and AI will be a key enabler.



What is the future of AI in fintech?

| Nitin Mishra, ED Tech, NPCI

I believe humanity evolves with technology. The future will have two kinds of people: those who create AI, and those who use AI effectively. Foundational models will keep improving, but domain-specific small models and agentic systems will dominate fintech. AI will shape fraud detection, risk scoring, operational efficiency, product innovation, and, customer-centric design. My daughter, who is very young, is already becoming a natural prompt engineer. That is the future-AI fluency will become universal.



Your final thoughts on the future of AI?

| Nitin Chugh, Deputy MD, SBI

AI is an irreversible wave-one that will keep evolving. Everyone will use it, just at different levels. We're already exploring AGI, and with quantum computing and energy-efficient models, the next leap is coming sooner than we expect. Academia-industry collaborations will be critical. At SBI, we partner with IIT Bombay, IIT Madras and IIT Delhi on research, quantum and AI models. The goal is to bring homegrown IP into banking. We must remain careful, aware and open-minded. AI will not render us irrelevant-it will expand our capabilities. Our task is to use it wisely, responsibly and confidently.

Insight to Integrity

“Redefining Vigilance & Efficiency with Artificial Intelligence”



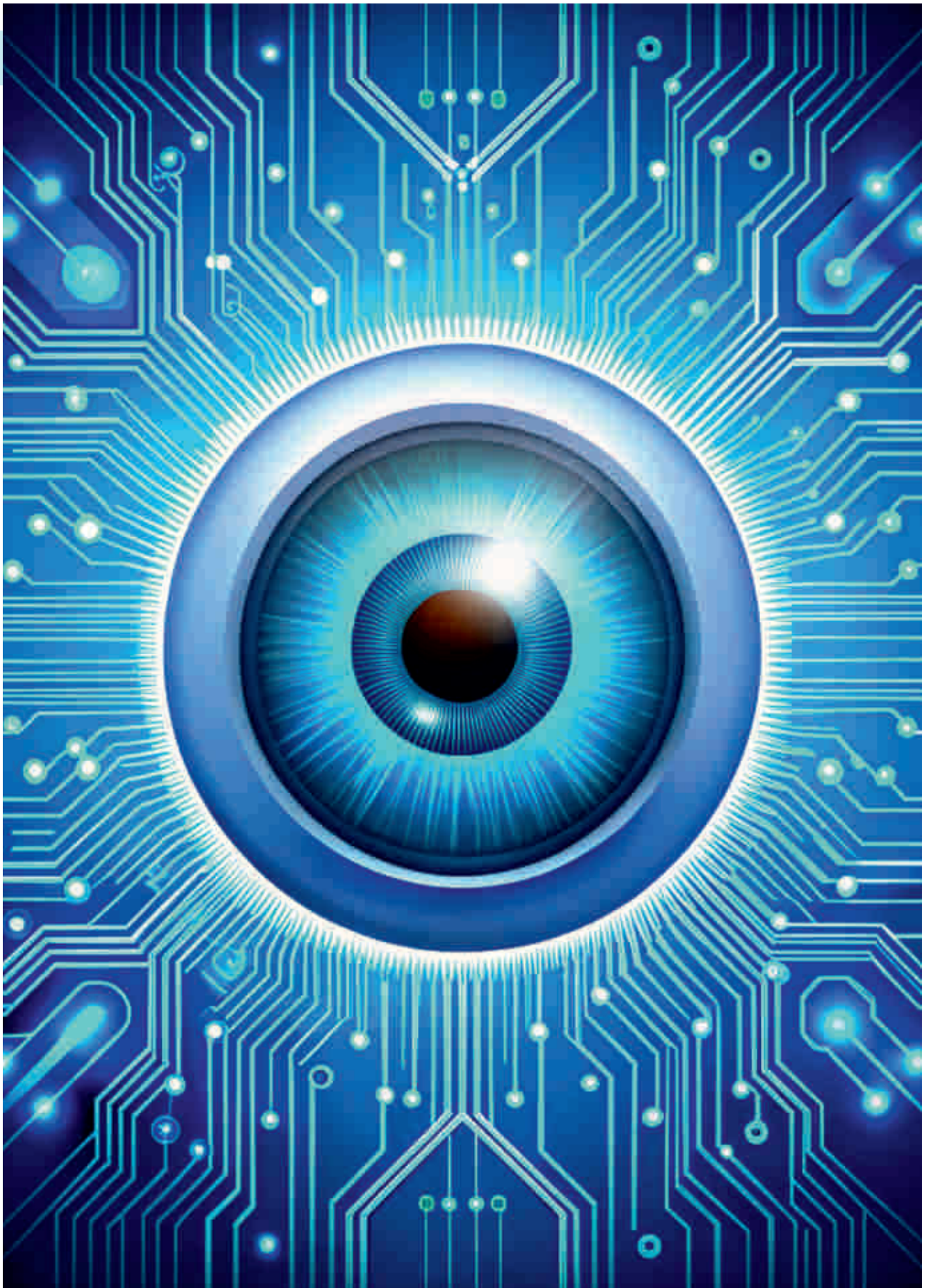
The panel brought together leaders from banking, payments, energy, automotive, railways and technology to discuss how AI, GenAI and Small Language Models (SLMs) are transforming efficiency, customer experience and preventive vigilance across India's largest institutions. SBI highlighted over 150 AI use cases, a patented high-risk branch detection model, and strong governance anchored in its Responsible AI Framework. NPCI showcased SLM-based dispute resolution, real-time fraud analytics and the need for velocity-driven vigilance in high-volume payments. Mahindra Group demonstrated enterprise-wide GenAI adoption-from email triage and diagnostics to connected-vehicle analytics-along with its ERS (Ethics-Responsibility-Security) approach.

IOCL emphasized predictive maintenance, pipeline

surveillance and the importance of building a unified data lake to scale AI responsibly. IRCTC illustrated AI success at national scale with multilingual IVR systems, bot detection, fraud mitigation and intelligent ticketing. IBM underscored the need to “Eliminate, Simplify, Automate” before deploying AI, and stressed governance, lifecycle management and quantum-AI convergence as the next frontier.

Across discussions, three themes emerged: *AI as a driver of operational excellence; preventive vigilance shifting from reactive audits to real-time, data-driven deterrence; and responsible AI-ethics, fairness, explainability and guardrails-becoming essential for trust.* Collectively, the panel affirmed that AI's future lies in scaled adoption, strong governance and continuous human-machine collaboration.







SECL

South Eastern Coalfields Limited

South Eastern Coalfields Limited

(A Govt. of India Undertaking)

PRODUCTION WITH PURITY DISPATCH WITH QUALITY



www.secl-cil.in



सतर्कता जागरूकता सप्ताह 2025

Vigilance Awareness Week 2025

सतर्कता : हमारी साझा जिम्मेदारी

Vigilance: Our Shared Responsibility



Central Vigilance Commission

Satarkata Bhawan, A Block, GPO Complex
INA, New Delhi-110023
www.cvc.gov.in