RESPONSIBLE BY DESIGN

Industry's Perspective on
India's AI Framework

PRIMUS PARTNERS®

August 2024

TABLE OF CONTENTS

LIST OF
TABLES & FIGURES

Foreword

Devroop Dhar
Co-Founder and Managing Director,
Primus Partners

Responsible Use of AI should be a Shared Global Priority

The rise of AI promises to transform the global economy and society. India, with its growing tech sector, is at the forefront of this change. At Primus Partners, we recognize the immense potential of Artificial Intelligence to drive growth, innovation, and improved quality of life. However, alongside this potential, we must also navigate the critical challenges of ethical use, accountability, and societal impact.

Our report, "Responsible by Design: Industry's Perspective on India's AI Framework," aims to provide a comprehensive overview of the current AI landscape in India. It addresses the pivotal principles of responsible AI, emphasizing the need for frameworks that prioritize ethics, transparency, and accountability. These principles transcend boundaries and the echoes shared globally. The recent regulatory stand by EU or developments in countries like US and Singapore have propelled necessary conversations for guiding frameworks. Our goal through the report is to suggest and foster an environment where AI can flourish responsibly, contributing to India's economic ambitions while safeguarding public interest.

As we journey towards becoming a trillion-dollar digital economy, it is imperative that we adopt a balanced approach to AI governance. This involves not only leveraging AI for its economic benefits but also ensuring robust regulatory frameworks that mitigate risks like deepfakes, and promote ethical standards. The findings and recommendations in this report reflect the collective take of industry leaders, and insights from subject-experts, offering a sustainable roadmap.

Primus Partners is committed to advancing the discourse on responsible AI. We believe that collaboration between the public and private sectors, guided by a shared vision, will be crucial in navigating the complexities of this transformative technology. Our report underscores the importance of continuous learning, capacity building, voluntary product iterations and responsible adoption, educating consumers, and proactive policy-making in shaping an AI landscape that is both innovative and responsible.

I extend my gratitude to all contributors and stakeholders who have provided their valuable perspectives and expertise in the preparation of this report. Together, we can ensure that AI serves as a force for good, driving inclusive growth and societal progress in India.

Executive Summary

The rapid evolution of artificial intelligence (AI) has the potential to significantly enhance India's digital economy, which is headed towards the $1 trillion mark. The exponential rise of the sector also presents immense opportunities for socio-economic progress. However, the expansion of AI brings challenges, particularly in ensuring responsible, safe, and accountable AI use.

This report is founded on an extensive study involving senior industry leaders and experts. The study aimed to capture a broad spectrum of views on the implementation of responsible AI in India. Through a structured survey as well as qualitative insights, the research explored key areas such as ethical AI governance, regulatory needs, operational challenges, and the impact of AI on decision-making and governance.

The study provides a nuanced understanding of the industry's readiness and challenges in adopting responsible AI practices. It highlights the need for frameworks that emphasize ethics, transparency, and accountability. The findings serve as a crucial resource for policymakers and industry stakeholders, offering actionable recommendations to foster a balanced and ethical AI ecosystem in India.

Key Findings from the Industry Survey

01

Impact on Decision-Making and Governance

3 out of 4 industry respondents believe that Responsible AI will enhance decision-making and governance by solving complex issues, accelerating transformation.

75%
of respondents believe that responsible AI will enhance decision-making and governance by solving complex issues and accelerating transformation.

22%
of respondents had a moderate to neutral view on this assertion.

Only 3%
disagreed with the statement, indicating broad support for responsible AI's positive impact.

02

Operational Factors for Ethical AI

Accelerating ethical AI ecosystem in India necessitates operational requisites like skill development and structural alignment in corporations.

73%
support establishing a dedicated AI Governance and Ethics Board within the corporate structure to oversee responsible AI practices.

Over 78%
emphasize the need for ongoing skill development and training to keep pace with AI advancements.

03
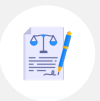
Perspectives on AI Regulation

89% of the industry believes that specific regulations are essential to ensure ethical practices, accountability, and transparency in AI development and deployment.

51%
of the surveyed industry leaders prefer a hybrid regulatory approach combining traditional frameworks with adaptive guidelines.

The industry favours...
a sensible, liberal, and market-friendly regulations to promote innovation while ensuring ethical standards.

04

Challenges in AI Development

Data Quality and Availability, Ethical Integration, and Ease in Regulatory Compliance are paramount in the development and deployment of AI tools in India.

76%
Data quality and availability are significant obstacles, with 76% highlighting issues such as data scarcity, biases in datasets, and complexities in data preparation and cleaning.

47%
believe that stringent and ad-hoc regulatory compliances remain a significant hurdle for the industry.

32%
of respondents cite the need for substantial computational resources and scalable infrastructure as critical challenges.

05

Sectoral Insights

Close to 70% of the industry strongly believes that more Responsible AI-based PPP engagements are required in critical sectors.

Public Sectors

AI is being increasingly implemented in public sectors such as healthcare and education, with initiatives like Bhashini and various Centres of Excellence driving innovation.

Major Indian IT Firms

and global firms in India are developing AI-powered platforms and training programs, indicating strong industry commitment to integrating AI across business operations.

06

Principle-Based Priorities

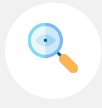Principle-based priorities shaping Responsible AI development in India include Privacy and Security, Transparency and Bias Mitigation.

Privacy and Security

61%

of the industry prioritize data privacy and security, stressing the importance of robust safeguards against data breaches and misuse.

Transparency and Explainability

53%

of respondents believe that AI systems should be designed to be transparent and explainable to ensure trust and accountability.

Fairness and Non-Discrimination

41%

emphasize the importance of AI systems being fair and non-discriminatory, ensuring equal treatment across different user groups.

Accountability

39%

highlight the need for clear accountability mechanisms for AI decisions, including audits, oversight and responsibility for outcomes.

Human-Centric Design

About 15%

support designing AI systems that prioritize human values and societal benefits, ensuring that AI serves the broader good.

On the regulatory front, the report discusses key developments include the possibility of a Digital India Act or equivalent measures towards revising the IT Act, which will address AI-related user harm, cybersecurity, algorithmic accountability, and transparency. The document also discusses integration of AI with competition law, citing the anticipated Digital Competition Act, which could encompass the scope of AI-driven market practices. The Economic Advisory Council to the

PM advocating for a dedicated 'agile' AI regulator is also highlighted. Furthermore, it highlights the implications of the Digital Personal Data Protection Act on AI, ensuring transparency, consent, and user rights in AI operations. Global AI regulatory perspectives from the EU, USA, China, Brazil, Singapore, Australia, and the UK are also reviewed, providing insights for India's AI governance. Three possible policy approaches that India can undertake are suggested – based on risks, sector and the underlying technology used.

Subsequently, the report outlines possible enforcement mechanisms for AI regulations or frameworks in India, emphasizing the roles of various bodies such as the Data Protection Board, and the Competition Commission of India. It proposes the formation of an RAI Accountability Committee to coordinate AI governance, an AI Technology Audit and Certification like Body, and suggests systemic voluntary responsible AI practices for companies. Importance of R&D and collaborative intelligence are also embedded in the report
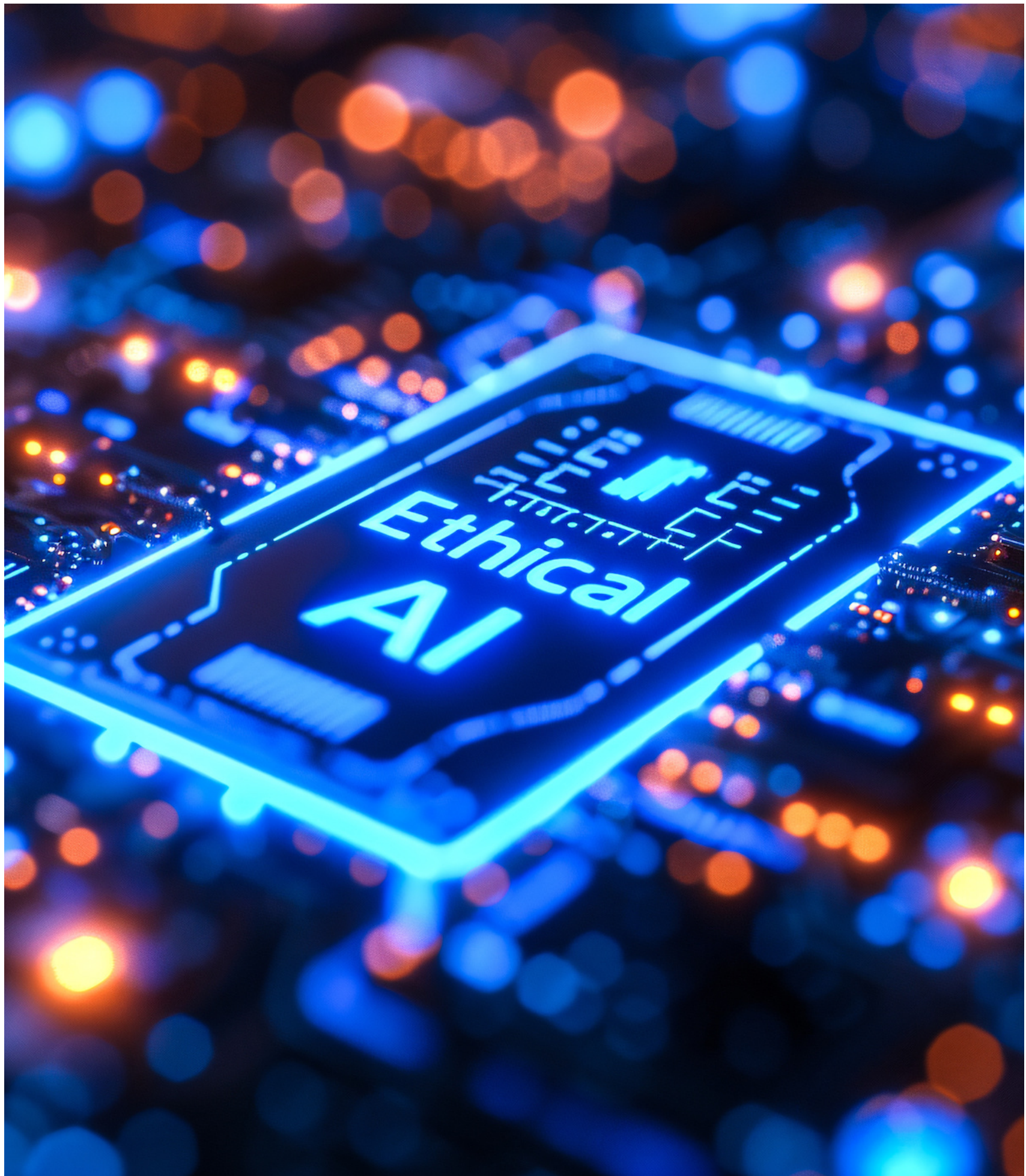
Overview of India's AI Landscape

India's digital economy has witnessed remarkable growth. This includes the tremendous opportunities in Artificial Intelligence.

AI is expected to add USD 967 Bn to the Indian economy by 2035 and USD 450−500 billion to India's GDP by 2025[1], accounting for 10% of the country's USD 5 trillion GDP target.

However, with this growth come complex challenges, especially in the realm of responsibility, safety, and accountability. The need for efficient and fair AI mechanisms has become increasingly evident.

With the overwhelming developments in the past one year globally, including Biden Administration's Executive Order on AI, OECD's AI Principles, G20 AI Guidelines, the EU's Artificial Intelligence Act, the G7 Hiroshima Principles on AI, the UNESCO Recommendations on AI ethics, GPAI's Declaration and UK's AI Safety Summit, India's current stand on prioritising innovation over compliance heavy regulation is plausible for various reasons. Firstly, India's AI industry is thriving with development of new Large Language Models (LLMs) and providing innovative solutions; this trend should continue for some time to scope how the sector further develops. This is also important so that it can keep pace with global developments. Secondly, India should effectuate discussions on R&D and growth of AI, for which supporting and testing innovation is the need of the hour. Thirdly, emerging innovations in AI also boosts India's Digital India campaign, making it progress towards a 'digitally empowered society and knowledge economy'.

India's current progress is also commendable. It ranks 1st in AI Skill Penetration as per the Stanford AI Index report 2023.[2] India's National Artificial Intelligence Portal, 'INDIAai' has also been launched as a one-stop digital platform for AI-related developments in India.
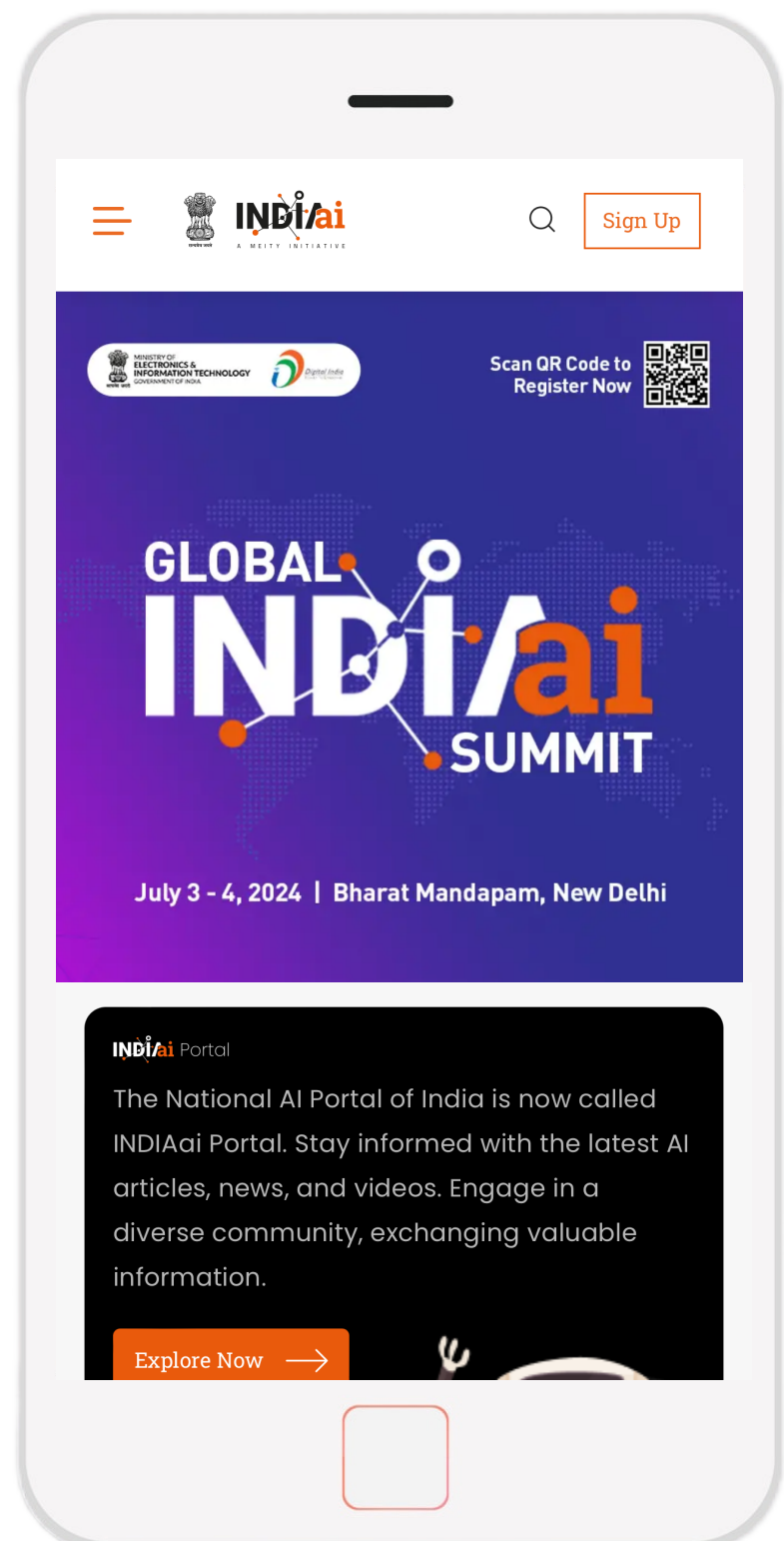
IndiaAI's several working groups, under Ministry of Electronics and IT's (MeitY) supervision, is collectively brainstorming on the vision and objectives of AI in India. The groups released its first set of recommendations in the form of a consolidated report[3]. The Ministry also drove India's Chair of the Global Partnership on Artificial Intelligence (GPAI) in December 2023.

_____

[1] https://www.pib.gov.in/PressReleasePage.aspx?PRID=1877503
[2] https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
[3] https://www.meity.gov.in/writereaddata/files/IndiaAI-Expert-Group-Report-First-Edition.pdf

Following up to GPAI's success, India also hosted another Global IndiaAI Summit during mid-2024 and another Global AI Summit by Kerala State government.[4] In addition, India is increasingly implementing AI applications in the public sector. This includes the introduction of Bhashini[5], an interactive platform designed under the National Language Translation Mission[6] of the government. It seeks to allow citizens to use the internet and digital services in their native languages. Academically as well, India has attracted global attention on the publications which have created significant impact.[7]

The Indian government in early 2024, also sanctioned a National IndiaAI Mission, allocating over Rs. 10,000 crore to foster the country's AI ecosystem growth. This was further supplemented by an additional Rs. 550 crore for FY2024-25.[8] This initiative aims to enhance computing access, data quality, and local AI skills, as well as foster industry partnerships, support startups financially, undertake socially beneficial AI projects, and promote ethical AI practices.

India has also rightfully made several packet developments across sectors in AI, including healthcare, climate, education, agriculture, BFSI, etc. In the corporate ecosphere, major Indian IT companies and global companies in India are developing platforms and services that leverage generative AI technologies. These initiatives range from AI-powered platforms for business solutions to extensive training programs for employees, indicating a strong commitment to integrating AI across business operations. However, there are looming challenges in the development and deployment of AI tools in India.

What is Responsible AI (RAI)?

While there has been no globally accepted norm or technical definition of what Responsible AI entails to, it however broadly represents a combination of value driven characteristics in parallel to the conventional understanding of AI; that is, an AI tool or practice which is trustworthy, designed with power dynamics and ethics in mind, and with minimized risk.

The Alan Turing Institute defines AI ethics, principally synonymous to RAI, as a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies.

Worldwide, including in India, governments are increasingly emphasizing the importance of integrating values and guidelines related to AI governance and ethics. Beyond just establishing these principles, it's crucial for India to devise ways to apply these standards across various sectors like public, private, and academic, ensuring a balance between fostering innovation and managing the potential risks.

[4] https://indiaai.gov.in/news/kerala-to-host-global-ai-summit-in-2024
[5] https://bhashini.gov.in/
[6] https://www.meity.gov.in/national-language-translation-mission
[7] https://www.nasscom.in/system/files/publication/AI-Patent-driving-emergence-of-india-Report_0.pdf
[8] https://www.indiabudget.gov.in/doc/eb/sbe27.pdf

Responsible AI Instrumental in Shaping India's Digital Economy:
Comprehensive Findings of the Study

The study was conducted with senior industry leaders and sectoral experts. Key findings have been summarized with visual representation for better understanding and further discussion.

Claims on AI are not often backed with good evidence. Data collected by Primus on the application of AI and its regulation provides some interesting insights. It is reassuring to see that captains of the industry are among the cheerleaders for making AI more responsible.

Dr. Deepak Mishra
Director and Chief Executive, ICRIER

Respondent / Surveyed Population

The respondents ranged from **Mid- to Senior-Level Industry Professionals.**

Technology and Digital Services Consulting

41.56%

Technology company which develops and innovates AI-based solutions
(or where AI is used in its service or product offers)

35.06%

Multiple

(Banking, Animation, Capital Markets, Auto Industry, Healthcare, Semiconductor, Retail, Industry Body)

15.58%

AI Startup
(< 5 years+)

3.9%

Core Responsible AI companies

3.9%

Fig 1 | Breakdown of Surveyed Population by sub-sector

01

An Overwhelming Majority feels that Responsible AI will Enhance Decision-Making and Governance by Solving Complex Issues, Accelerating Transformation

Will Responsible AI enhance decision-making and governance?

Moderately agree to neutral
22%

Strongly agree
75%

3%
Disagree
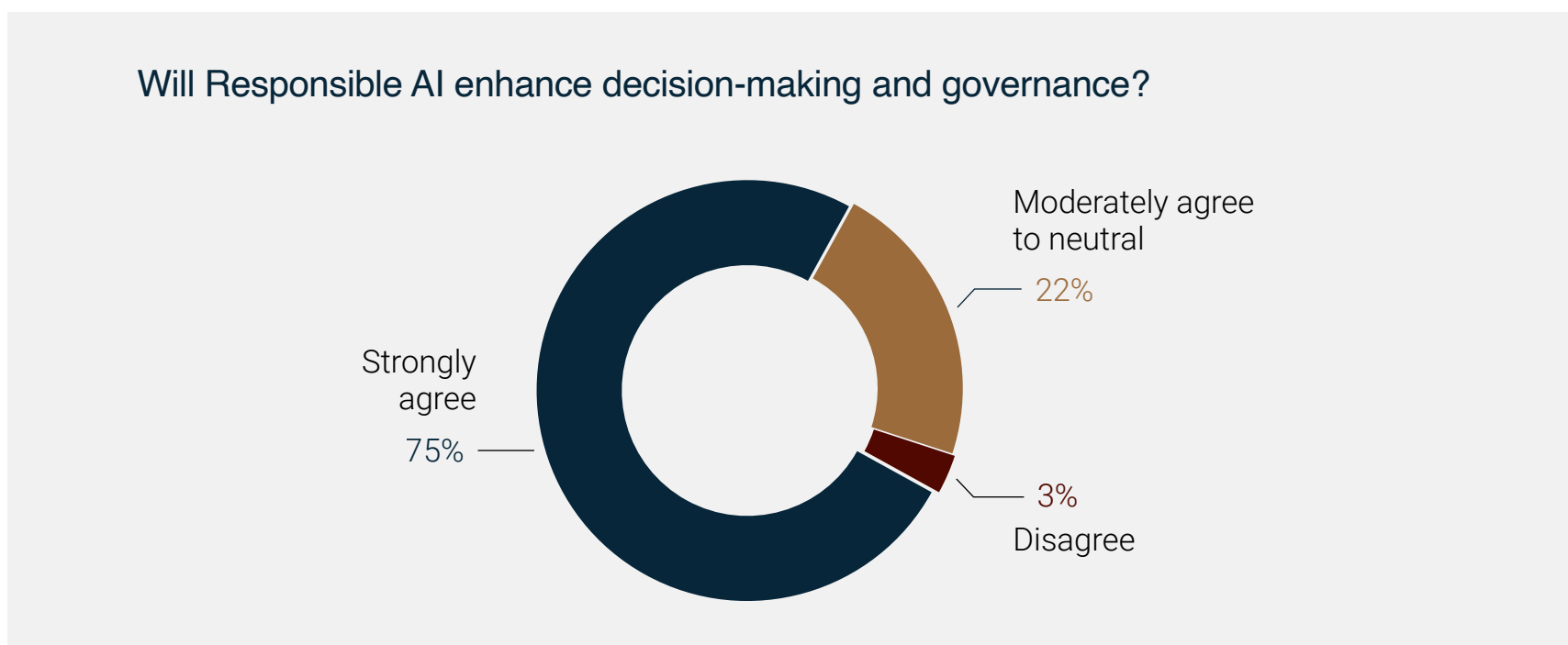
Fig. 2 | Will Responsible AI enhance decision-making & governance?

The findings highlight the importance of integrating ethical and responsible AI frameworks within organizational strategies. By leveraging AI's capabilities in a responsible manner, organizations can achieve more informed and effective decision-making processes, ultimately leading to improved governance and accelerated transformation across various sectors.

This goes for both corporations and government mechanisms. However, it is advisable that only simpler AI tools or models should be used when it comes to critical decision-making. This will help ensure that decisions can be explained more easily to end users.

Addressing Ethical Challenges from AI-Based Outcomes

Responsible practice of AI will help address challenges and issues that occur in the use of traditional artificial intelligence models and deployment. The following are some of the prominent concerns both from India and globally that AI leveraging organisations are facing, which is impacting their decision making ability and reputation:

Sensitive and Biased Outcome

Outcomes which are biased or misdirected, esp. from Generative AI models have been an issue for both governments and industry alike. There have been various instances of these models being biased and inaccurate[9]. Reports have shown that biased recommendations made by AI systems can adversely impact human decisions in professional fields. Moreover, they also show that such biased recommendations can influence human behaviour in the long term[10].

Privacy Concerns

Use of data and large datasets is a conundrum. While AI models, often also regarded as advanced statistical models integrated with code or predictive models derived from probability distribution[11], are made efficient with the training of data fed into it, the same can also have privacy infringement concern. Copyright infringement is another essential extension of violation of information and copyright, especially for large corporates and publishers. India's Data Protection Law provides certain level of security against the use of one's personal data, but more focused interventions are requisite to address the question of 'Can AI be by choice?'.

Flawed Algorithms

Flawed or substandard algorithms can be an outcome of various reasons. Lack of adequate technical expertise and resources, lack of awareness amongst the technologists, lack of oversight, or lack of certified data fed.

AI Hallucinations

Refers to the phenomenon where AI models, particularly generative AI models like language or image generators, produce outputs that are factually incorrect, nonsensical, or disconnected from reality. For example, generative AI models generating URLs, code libraries, and even people that do not exist, along with made-up news articles, books, and research papers. This can be particularly detrimental in research or legal contexts, where accuracy is paramount, and becomes increasingly relevant as reliance on AI systems for decision-making and information dissemination grows.

---

[9] https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/ and https://www.prolific.com/resources/shocking-ai-bias
[10] https://www.nature.com/articles/s41598-023-42384-8#:~:text=Our%20results%20show%20that%20biased,behaviour%20in%20the%20long%20term.
[11] https://www.youtube.com/watch?v=5t1vTLU7s40&pp=ygUWeWFubiBsZWN1biBsZXggZnJpZG1hbg%3D%3D

Deepfakes

India ranked first in the potential risks of misinformation and disinformation per the World Economic Forum Executive Opinion Survey 2023[12]. The most discussed of all the concerns from AI's perspective are Deepfakes. In simple words, deepfakes are videos or images creating delusion with the use of deep learning, AI and photoshopping techniques to spread disinformation or for unethical commercial gains. India's 2024 General elections is the most prominent example here[13]. The technologies namely, GANs (Generative Adversarial Networks) or ML (Machine Learning) are interplayed to create such videos or images. It also has certain positives, such as in medical training, VR applications and simulations, journalism.

Accountability

The accountability quotient is important to answer the following question – in case of any wrongdoing or bias in the outcome of AI models, is it the model which is to be made accountable or the creator of the model or the source company? As an example, should we consider a chatbot an intermediary or a publisher?

Failure to mitigate these challenges can lead to severe consequences for organisations and corporations, which include:

1. Reputation Damage

Companies can suffer significant reputational harm if AI systems produce biased, inaccurate, or harmful outcomes. This can lead to a loss of trust among consumers and stakeholders.

2. Legal Repercussions

Non-compliance with privacy laws and regulations, such as India's Data Protection Law, can result in legal actions, fines, and sanctions against organizations.

3. Financial Losses

Biased or flawed AI outcomes can lead to poor decision-making, resulting in financial losses and decreased profitability.

4. Loss of Consumer Trust

Privacy infringements and data misuse can erode consumer confidence, leading to a decline in customer base and loyalty.

5. Operational Disruptions

The spread of misinformation through deepfakes and other AI-generated content can disrupt operations and communication within and outside the organization.

6. Ethical and Social Backlash

Companies may face ethical scrutiny and social backlash for deploying AI systems that perpetuate bias, violate privacy, or contribute to misinformation.

[12] https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf
[13] https://www.bbc.com/news/world-asia-india-68918330

Responsible development and deployment of AI systems encompass a wide range of concerns, ranging from data handling and privacy to performance and monitoring of deployed systems. Adoption of these principles throughout the AI pipeline would make it much easier for organizations to get stakeholder confidence and buy-in for using AI in decision-making.

Professor Balaraman Ravindran
Head - Centre for Responsible AI (CeRAI), IIT Madras;
Senior Member - Association for Advancement of AI (AAAI),
Washington D.C.

02

Data Quality and Availability, Ethical Integration, and Ease in Regulatory Compliance are paramount in the Development and Deployment of AI tools in India

We assessed some of the common concerns and challenges that the industry and startups face while in the process of developing AI models and during deployment of such AI tools. The assessment was on the grounds of data, on integrability, impact considerations on society's ethical values, computational resources, regulatory compliances and cross-border trade restrictions. The key aspects include the following: Addressing data quality and availability is fundamental, as high-quality data is crucial for the accuracy and reliability of AI models. Ethical integration ensures that AI systems operate fairly and transparently, maintaining public trust and adherence to societal values. Additionally, simplifying regulatory compliance is essential to foster innovation while protecting user interests. Fig.3 shows the detailed outcome of the assessment.

Data Quality and Availability
(also includes data scarcity, biases in datasets, and the complexities of data preparation and cleaning)
75.7%

Integration with Existing Systems and Processes
(also encompasses challenges related to compatibility, interoperability, and the transformation of business processes to leverage AI effectively)
56.8%

Ethical Considerations and Societal Impact
52.7%

Regulatory Compliances and Legal Issues
(Including ex-ante obligations)
47.3%

Computational Resources and Scalability
32.4%

Cross-Border Trade Restrictions
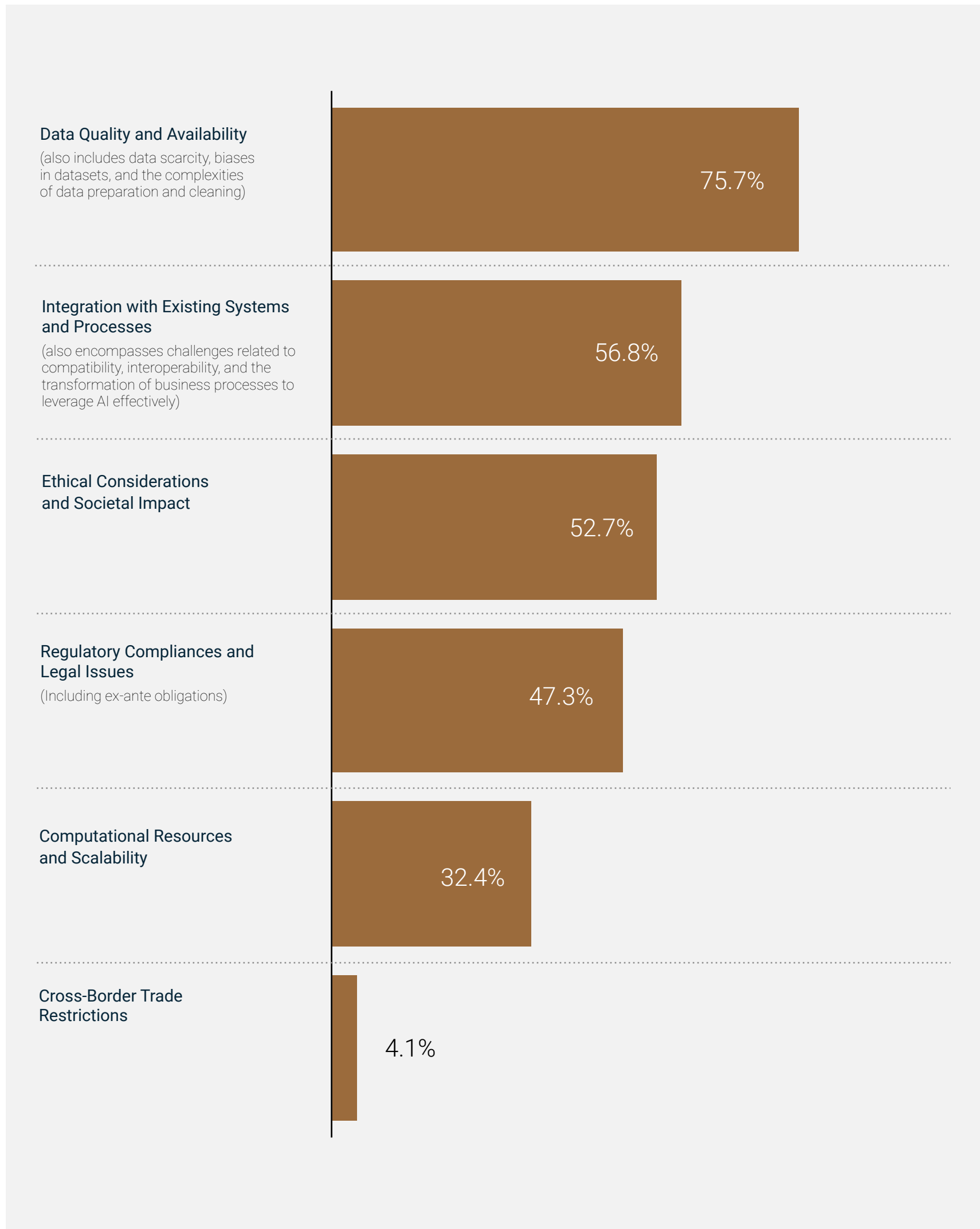**4.1%**

Fig. 3 | Principal Obstacles in AI Development, Deployment and Training

01

Data Quality and Availability

Current Scenario

One of the paramount challenges in AI development, which 75.7% of the respondents have agreed to, is the quality and availability of data. Issues such as data scarcity, biases in datasets, and the complexities of data preparation and cleaning are prevalent. These challenges are critical as they directly impact the performance and reliability of AI systems.

Recommendation

In the Indian context, a priority-based approach with first improving data quality and availability across critical sectors can significantly enhance AI applications. Establishing robust data governance frameworks and investing in advanced data infrastructure will ensure data integrity and efficient handling of large volumes of data. Promoting interoperable data standards will facilitate seamless data sharing, while innovative data collection techniques like IoT and remote sensing will enhance data quality. Additionally, investing in training programs for data management skills will significantly improve data availability and quality, thereby enhancing AI applications across sectors.

02

Ethical Integration

Current Scenario

52.7% industry professionals have argued for integrating ethical considerations into AI development, to ensure that AI systems operate fairly and transparently. Ethical concerns include addressing biases, ensuring accountability, and protecting user privacy.

Recommendation

With the rapid adoption of AI in India, there is an increasing need to embed ethical frameworks within AI systems to prevent unintended consequences such as job displacement, privacy violations, and biased decision-making. This aligns with global trends where ethical AI is becoming a focal point in regulatory discussions.

03

Ease in Regulatory Compliance

Current Scenario

Stringent and Ad-hoc regulatory compliances remain a significant hurdle. 47.3% believe that over-regulation and over-compliance can be a deterrent to AI's growth. The Indian government's approach to AI regulation has been oscillating between promoting innovation and imposing necessary safeguards to mitigate potential risks. The introduction of the Digital Personal Data Protection Act 2023 marks a step towards addressing data protection and privacy concerns, though it still has gaps that need addressing. The influx of multiple legislations and compliances can however create a burden and avoidable confusion for the industry.

Recommendation

Simplifying regulatory compliance can spur innovation while ensuring that AI technologies are deployed responsibly. It is crucial for India to develop a balanced regulatory framework that protects users without stifling innovation. This involves creating clear guidelines and standards that align with India's unique socio-economic context. Moreover, advisories and directives should encompass prior expert consultations before being notified.

04

Computational Resources and Scalability

Current Scenario

AI development demands substantial computational resources, including high-performance hardware and scalable infrastructure. The need for advanced GPUs, cloud computing services, and scalable data storage solutions is critical to support the intensive computational requirements of modern AI algorithms. 32.4% of the respondents say that the current systems are inadequate.

Recommendation

India's AI ecosystem is growing, but the availability of computational resources remains a bottleneck. Investments in infrastructure, such as the National AI Mission, which includes setting up AI research centers and supercomputing facilities, are essential to address these challenges.

05

Integration with Existing Systems and Processes

Current Scenario

56.8% of the industry respondents have voiced their concern on the difficulty in integrating growing AI models into the current tools, mechanisms and systems of business operations. Integrating AI with existing systems involves challenges related to compatibility, interoperability, and transforming business processes to leverage AI effectively. Organizations must ensure that AI solutions can seamlessly interact with legacy systems and modern IT infrastructure.

Recommendation

Successful integration of AI into business operations can drive efficiency and innovation. In India, sectors like finance, healthcare, and manufacturing are already seeing AI integration, but a broader and more coordinated effort is needed to overcome technical and organizational barriers .

06

Cross-Border Trade Restrictions

Current Scenario

Just above 4% of the respondents fear that cross-border trade restrictions can impact the global collaboration essential for AI research and development. These restrictions can potentially limit the access to international data sets, talent, and cutting-edge technologies.

Recommendation

For India to become a global AI hub, it must navigate such restrictions effectively. Strengthening international partnerships and advocating for open data exchange can help mitigate these barriers and promote a collaborative AI research environment.

03

Principle-based Priorities Shaping India's Responsible AI Development Include Privacy and Security, Transparency and Bias Mitigation

We assessed the industry on prioritising various principles and areas that are the foundational tenets of responsible AI deployment and use. These include Bias mitigation, Accountability, Privacy, Security, Explainability, Intellectual property, Algorithmic collusion and competition, as well as use of emerging technologies. Each of these principles plays a crucial role in shaping a responsible AI framework. Bias mitigation is essential to ensure fairness and prevent discrimination in AI systems, while accountability mechanisms ensure that there is a clear line of responsibility for AI outcomes. Privacy and security measures protect user data and maintain trust, and explainability ensures that AI decisions are transparent and understandable. Protecting intellectual property rights fosters innovation, and addressing algorithmic collusion and competition safeguards market fairness. Fig. 4 shows the detailed outcome of the study.

40.5%
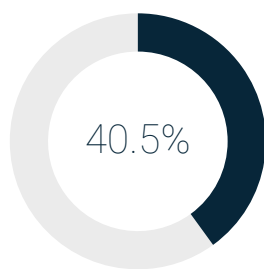
Bias Mitigation and Fairness

Mitigating biases and ensuring fairness, especially in AI algorithms. training and outcomes

The survey highlights that industry leaders consider mitigating biases and ensuring fairness in AI algorithms, training, and outcomes as a top priority. This focus is driven by the need to build trust and reliability in AI systems, ensuring equitable treatment and reducing discriminatory practices. However, more discussions are required on the 'subjectivity' of what bias encompasses. This becomes especially important on building the specifications of biases around India's diverse language use.

52.7%

Explainability and Transparency
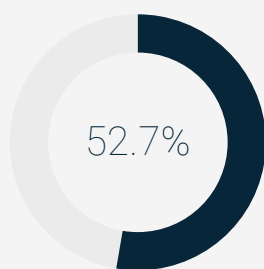
Ensuring AI systems are understandable and their decisions can be explained to users and stakeholders

According to the survey, respondents emphasize the importance of transparency in AI systems. Ensuring that AI operations are understandable and their decision-making processes, including criteria used, are explainable to users and stakeholders is crucial for fostering trust and accountability. A credible way forward can be institutionalising transparency and accountability in AI systems. This involves mandating 'data cards' or model cards (similar to nutrition levels in food). Transparency can also be on the governance structures of companies, and around data sharing (with reference to purpose limitation from DPDP Act 2023). Algorithmic transparency also becomes important to reduce impact on digital creators.

60.8%
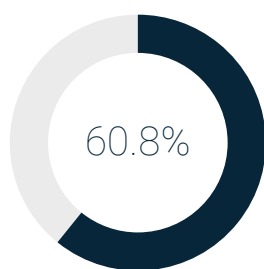
Privacy and Security
(including deepfakes)

Protecting user privacy and securing data, including mitigating the risks posed by deepfakes

The survey results also show that industry professionals prioritize privacy and security in AI development. This includes addressing the challenges posed by **deepfakes**, which have significant implications for misinformation and user privacy. Strong measures are required to protect personal data and maintain the integrity of AI-generated content. Established tech platforms can help government integrate deepfake monitoring with their current mechanisms of addressing misinformation. As an example, initiatives like 'PIB Fact Check' or 'DigiKavach' can encompass a deepfake alert infrastructure which can be built with advanced tech tools. Geopolitically, deepfakes also need to be looked at from a cross-country liability chain intervention instead of national silos.

25.7%
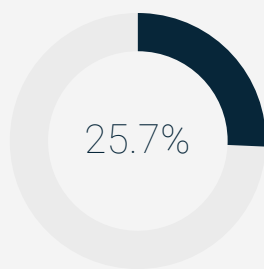
Use of Emerging Tech
(like blockchain)

Leveraging emerging tech, such as blockchain, to enhance AI systems' capabilities and trustworthiness

The industry survey further reveals that about 26% of participants see the integration of emerging technologies, such as blockchain, as essential for advancing AI capabilities. Emerging and future technologies hold immense possibilities, and can enhance the transparency, security, management and traceability of AI systems, thereby supporting more robust and trustworthy AI applications.

39.2%

Accountability

Establishing clear accountability mechanisms for AI system outcomes and decisions

21.6%

Safeguarding IPR

Protecting intellectual property rights within AI development and deployment

21.6%

Controlling Algorithmic Collusion

Preventing & controlling collusion between algorithms that can lead to anti-competitive practices

6.8%

Upholding Watermarked AI Content

Ensuring AI-generated content is appropriately watermarked for authenticity and traceability

14.9%

Assessing Socio-Cultural Impact

Evaluating the socio-cultural implications of AI applications to ensure they are beneficial to society

Data Cards are structured summaries of essential facts about various aspects of ML datasets needed by stakeholders across a dataset's lifecycle for responsible AI development.

Fig. 4 | Key Elements Shaping Responsible AI Development in India

With the growing use of AI, It is time to operationalise the key principles of AI Governance at scale. As a first step, it is necessary for AI developers and deployers to transparently communicate their practices on responsible AI and demonstrate how the practices are put into effect. This is essential where automated decision making can impact user rights or exacerbate harms. Governments and industry need to co-develop and together evolve good practices on responsible development and use of AI.

Ashish Aggarwal
VP and Head of Public Policy, NASSCOM

04

Accelerating Ethical AI Ecosystem in India Necessitates Operational Requisites like Skill Development and Structural Realignment

Industry leaders were asked about the operational requirements that are most essential towards accelerating India's ethical AI ecosystem. These operational requirements are viewed as critical for establishing a foundation where ethical AI practices can flourish and be sustained across various sectors in India. The factors put forward included a dedicated Board or Oversight mechanism, Putting theory to practice, Skill development and training, Multi-stakeholder engagements that can act as catalysts. Fig. 5 shows the outcome of the study. A map for voluntary adoption of responsible practice by organisations is also subsequently suggested, while also upholding R&D and consumer awareness in the process.
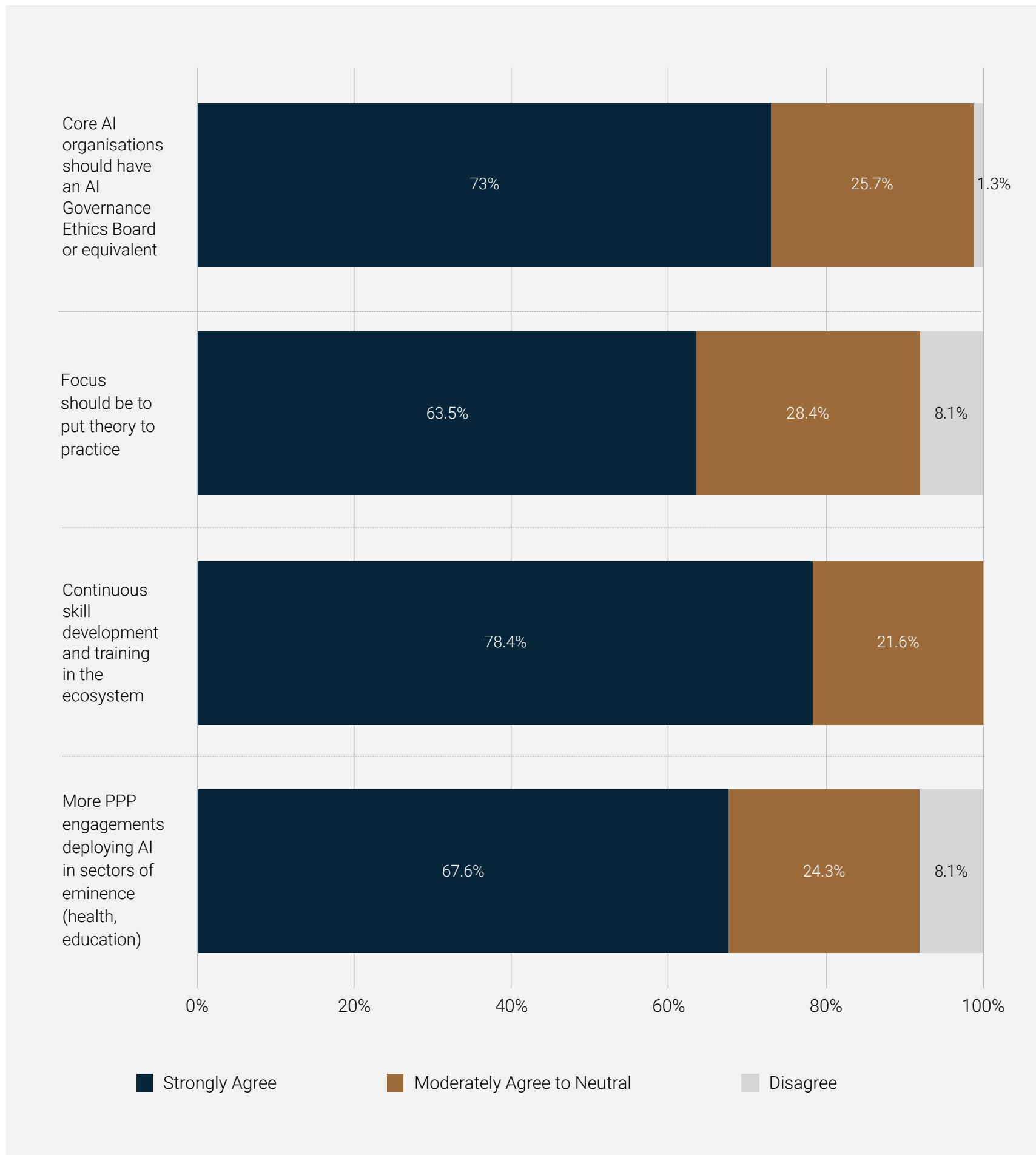
Fig 5 | Operational Factors crucial for accelerating Ethical AI Ecosystem in India

73%

Most respondents strongly agreed on the necessity of these factors, with 73% supporting the establishment of a dedicated and focused AI ethics board within the corporate governance structure.

This involves the leadership who continually work to ensure that AI systems are functioning properly, safely, and responsibly. When an AI system for credit approval, for example, is found to be discriminating against certain groups (as has happened before), ethics managers or the 'Ethics Board' are responsible for investigating and resolving the issue.

Similarly, data compliance officers ensure that the data used by AI systems complies with GDPR and other consumer-protection regulations. Another related example involves ensuring responsible management of information by AIs.

Like many tech companies, Apple uses AI to gather personal details about users interacting with its devices and software. This aims to enhance the user experience, but unrestricted data collection can compromise privacy, anger customers, and violate laws. Apple's "differential privacy team" ensures that while the AI learns as much as possible about a user group statistically, it protects individual users' privacy.

78.4%

Also, a significant 78.4% of respondents emphasize the need for ongoing skill development and education in the AI sector.

India faces a substantial AI skill gap, with a current demand for AI professionals far exceeding the supply. As of 2023, India had 4 lakh AI professionals but needs approximately 6.5 lakhs to meet industry demands, which are further expected to reach 1 million by 2026[14]. This gap highlights the necessity for collaborative efforts between the government, academia, and the corporate sector to design comprehensive skill development programs and promote continuous education in emerging technologies like AI. Also, India has the largest reserve of young and skilled tech talent, with an estimated 18 million STEM graduates by 2027. The adequate and proportional channelisation towards AI should be a high priority within STEM.

67.6%

of the industry leaders believe that there should be more PPP agreements in the subject around critical sectors. However, it is important to note that AI should be used as a tool for ameliorating these critical sectors rather than making direct decisions that potentially govern human lives.

63.5%

Putting theory to practice is also an industry priority, with 63.5% strongly agreeing with the application side of responsible practices. This can be attained by developers adopting voluntary mechanisms, and inculcating ethos of 'Secure and Responsible by Design'.

---

[14] https://wheebox.com/assets/pdf/ISR_Report_2024.pdf ; and https://www.weforum.org/agenda/2024/01/to-truly-harness-ai-we-must-close-the-ai-skills-gap/#:~:text=URL%3A%20https%3A%2F%2Fwww.weforum.org%2Fagenda%2F2024%2F01%2Fto

Voluntary Adoption of Responsible Practice by Companies & Organisations Building or Using AI: 'Secure and Responsible by Design'

Responsible AI enables trust in AI native or AI integrated systems to give confidence that such AIs are acceptable by the society and delivers value to its purpose. However, at times, companies building or using AI sometimes unsure how to achieve responsible AI and may decide not to harness its fair potential. To address this, companies can use tools and methods such as SHAP and LIME that make AI decisions more understandable, Microsoft's Fairlearn and Google's Model Card Toolkit etc to check for and reduce bias, and ethical guidelines from the IEEE and the European Commission provide valuable standards

Rohit Purohit
Chief Executive Officer, ViitorCloud Technologies

Technologists can apply the four different stages while ideating and subsequent adoption on an AI model:

Intent	Identify	Evaluate	Plan
The intent behind the work and investment on AI model	Identify the source of data and how it is being collected	Evaluate the data sources and impact of proposed AI model	Planning of roll-out including deployment & (business) execution
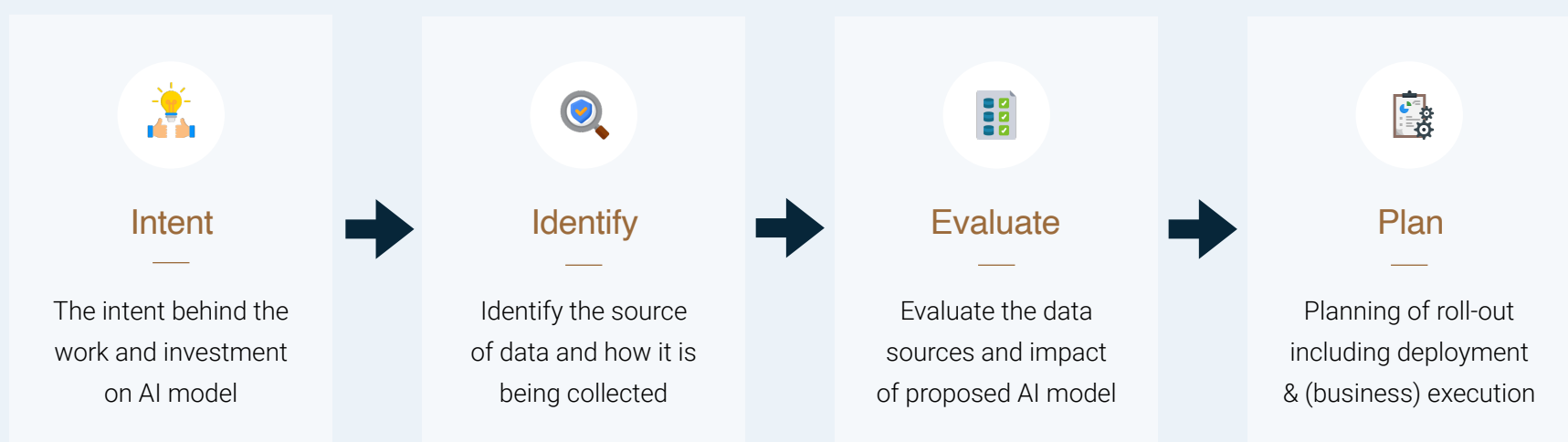
Evaluation Phase: Accommodating the Approach of "Tech Ethics by Design" vis-à-vis Ethical Standards Application

The following entails an outcome-based evaluation approach and thereon, putting necessary guardrails wherever required.

Layers of Impact

Classifying and laying out the impact of the AI model or tool, which includes outcomes that **are intended and known,** and secondly, **unintended and possibly known impact.**

Ethical Impact Assessment

On the second classification of supposed impact ideated, that is, unintended and possibly known, a deeper impact assessment to be done by further dividing them into **Beneficial** (or Benefits) and **Harmful** (Harms).

Ethical Hacking

✦ Firstly, understanding what the organisation's values and principles are on AI and overall tech deployment.

✦ Mapping out the rights of the individuals and end users or digital consumers.

✦ Once the organizational values and rights of the users are well established, this third step is an attempt to mitigate the potential harms that would be mapped from Step 2; Accordingly thus, requisite **moderations and changes in the Design and building of the model** is to take place. This also includes ensuring and keeping in consideration the AI system remains in **compliance with evolving legal and regulatory standards,** as well as internal **(corporate) governance** policies.

Prioritising specialised focus and need for R&D

India's journey in the realm of Responsible AI innovation, particularly in terms of research and development (R&D), reflects a complex interplay of ambition, potential, and the imperative need for specialized focus. The 2019 allocation[15] of ₹7,000 crore (approximately $1 billion) for NITI Aayog's AI program or the allocation in 2024 under India AI Mission, are significant steps, however still behind compared to global giants like the US and China.
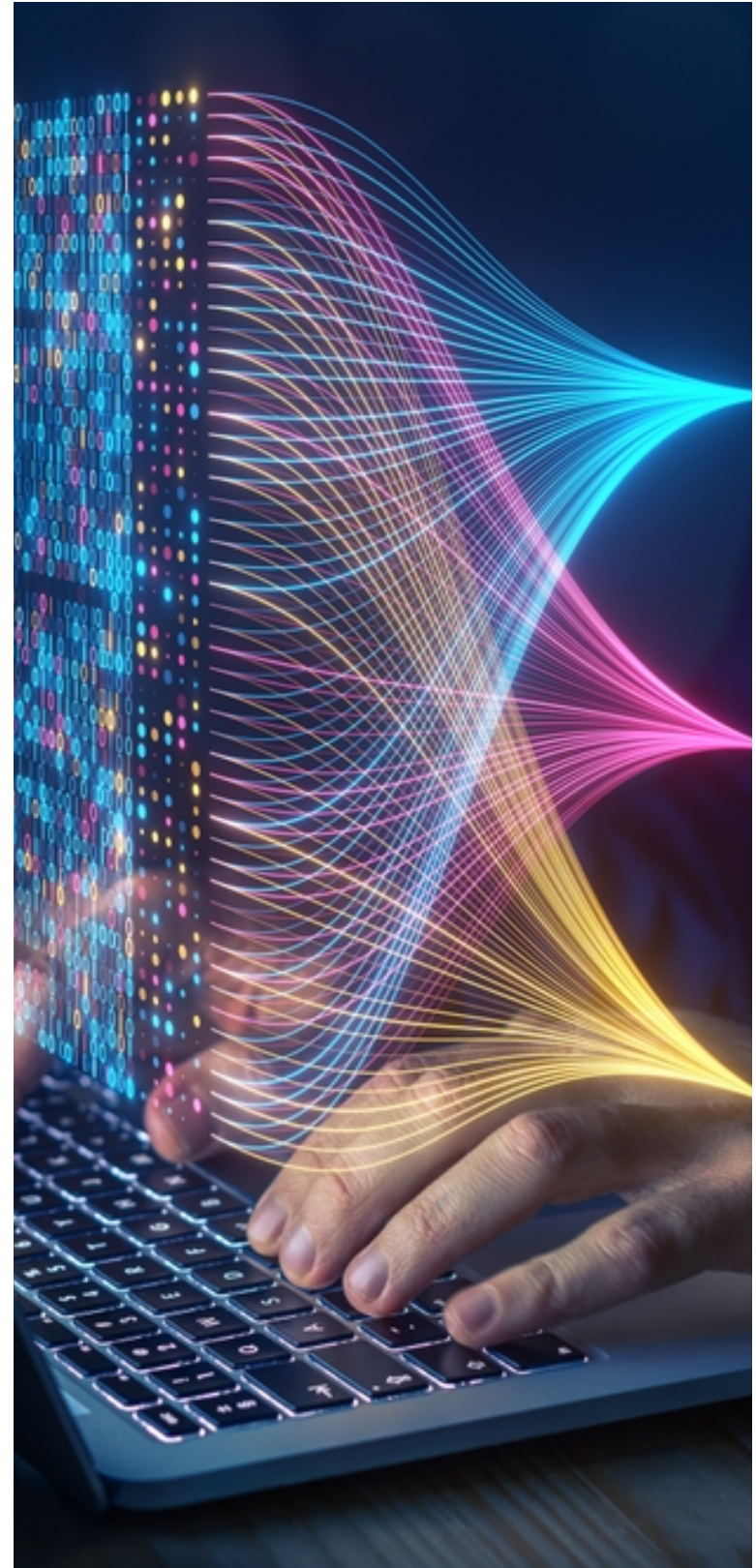
India's standing in **AI patent filings** is noteworthy. The country ranked 8th in the top 10 countries for AI patents[16], mainly in electronics and healthcare industries. This growth demonstrates India's emerging capability in AI innovation, yet there remains a substantial gap compared to the global leaders . The deployment of 5G infrastructure and recent AI initiatives are set to further catalyse this growth, providing the necessary technological backbone for advanced AI applications across various sectors, including healthcare, agriculture, and education.

In terms of **institutional capacity**, India has made some strides. The establishment of Centres of Excellence[17] (CoEs) and the Centre for Research on AI (CeRAI) are steps in the right direction. These institutions play a crucial role in fostering an ecosystem conducive to AI innovation, offering resources, expertise, and a platform for collaboration between academia, industry, and government. However, to truly leverage these institutions, there needs to be a concerted effort towards cohesive a policy framework, ethical standards for AI, aggressive private investments, skilling and continuous reskilling of the talent pool, and a sustainable roadmap for AI development in India.[18]

[15] https://www.moneycontrol.com/news/india/niti-aayog-gets-rs-7000-cr-for-artificial-intelligence-project-report-4419411.html
[16] https://nasscom.in/knowledge-center/publications/ai-patents-driving-emergence-india-ai-innovation-hub#:~:text=Key%20Findings&text=Innovation%20in%20AI%20has%20gained,terms%20of%20AI%20research%20papers
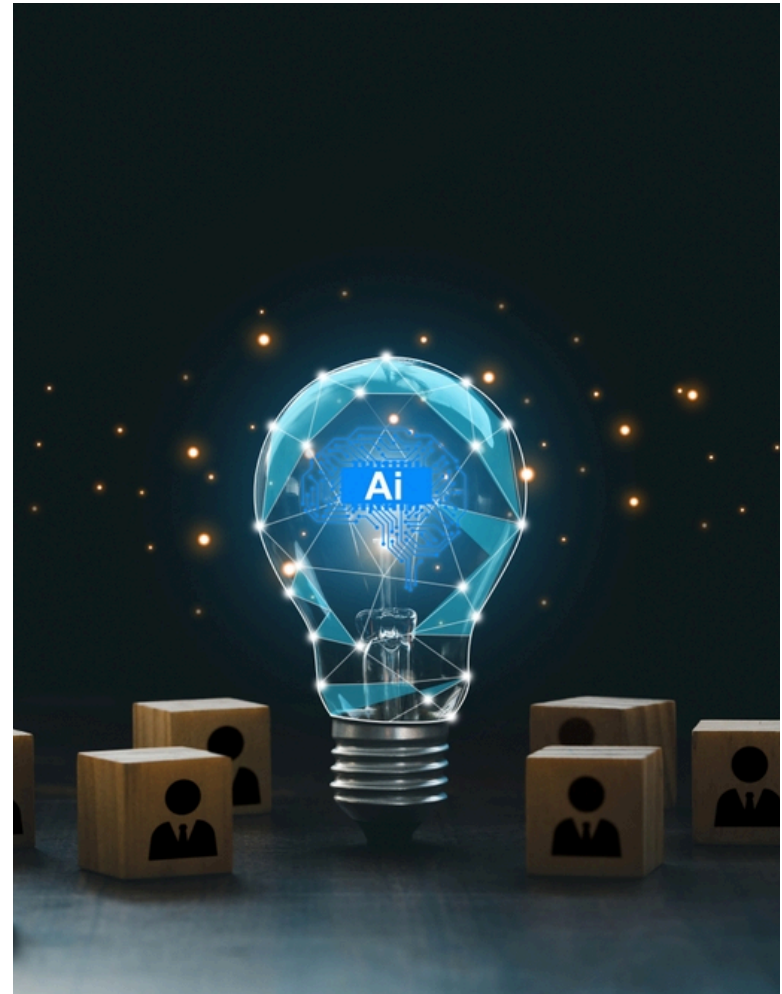[17] https://www.indiabudget.gov.in/doc/Budget_Speech.pdf
[18] https://cerai.iitm.ac.in/

Empowering Users: Creating Awareness Among Consumers on Safe and Responsible AI Use

To effectively raise awareness about the safe and responsible use of AI in India, a multi-faceted approach is advisable. This would include educational campaigns and workshops across the country, integrating AI education into school and university curricula, and initiating community engagement programs for interactive learning. A dedicated 'AI Consumer and Education Program', as a joint inter-departmental activity can be a start. Public-private partnerships could be leveraged to develop and disseminate educational materials. Developing online platforms and mobile applications can facilitate access to information, while media engagement can help in reaching a wider audience. Utilizing India's digital infrastructure for awareness programs, establishing feedback mechanisms for consumer experiences, promoting ethical AI guidelines and standards which prioritises user rights, are also crucial steps in this direction. The recent Dark Patterns Guidelines[19] by the Department of Consumer Affairs also translate well for AI models which have active consumer interaction.

---

[19] https://consumeraffairs.nic.in/sites/default/files/The%20Guidelines%20for%20Prevention%20and%20Regulation%20of%20Dark%20Patterns%2C%202023.pdf

Consumers deserve AI that prioritizes their rights and enhances their lives without compromising privacy or fairness. Responsible AI must include transparency in the use, accountability for outcomes, and ethical guidelines to protect users from biases and ensure that their data is handled with the utmost care. By fostering consumer awareness and implementing robust standards, we can build a digital future that is both innovative and secure.

Dr. Fauzia Khan
Member of Parliament, Rajya Sabha

05

The Industry has a Positive Outlook Towards Regulation that is Progressive, Liberal and Market friendly

The industry survey reveals a positive perceived outlook towards regulations, with the prerogative that these governing rules are drafted sensibly, are liberal, and market-friendly. This sentiment indicates a preference for regulatory frameworks that foster innovation and growth while ensuring fair competition and consumer protection.

Moreover, the industry realizes that a well-defined regulatory path will provide a clear roadmap and clarity of thought, which are essential for strategic planning and decision-making. By offering guidance and predictable rules, such regulations can help businesses navigate the evolving market landscape more effectively, reducing uncertainty and fostering long-term growth.

The Requirement for Regulatory Framework

● **Regulatory Frameworks needed:**
Specific regulations are essential to ensure ethical practices, accountability & transparency in AI development and deployment.

● **No Specific Regulations required currently**

89.2%

10.8%
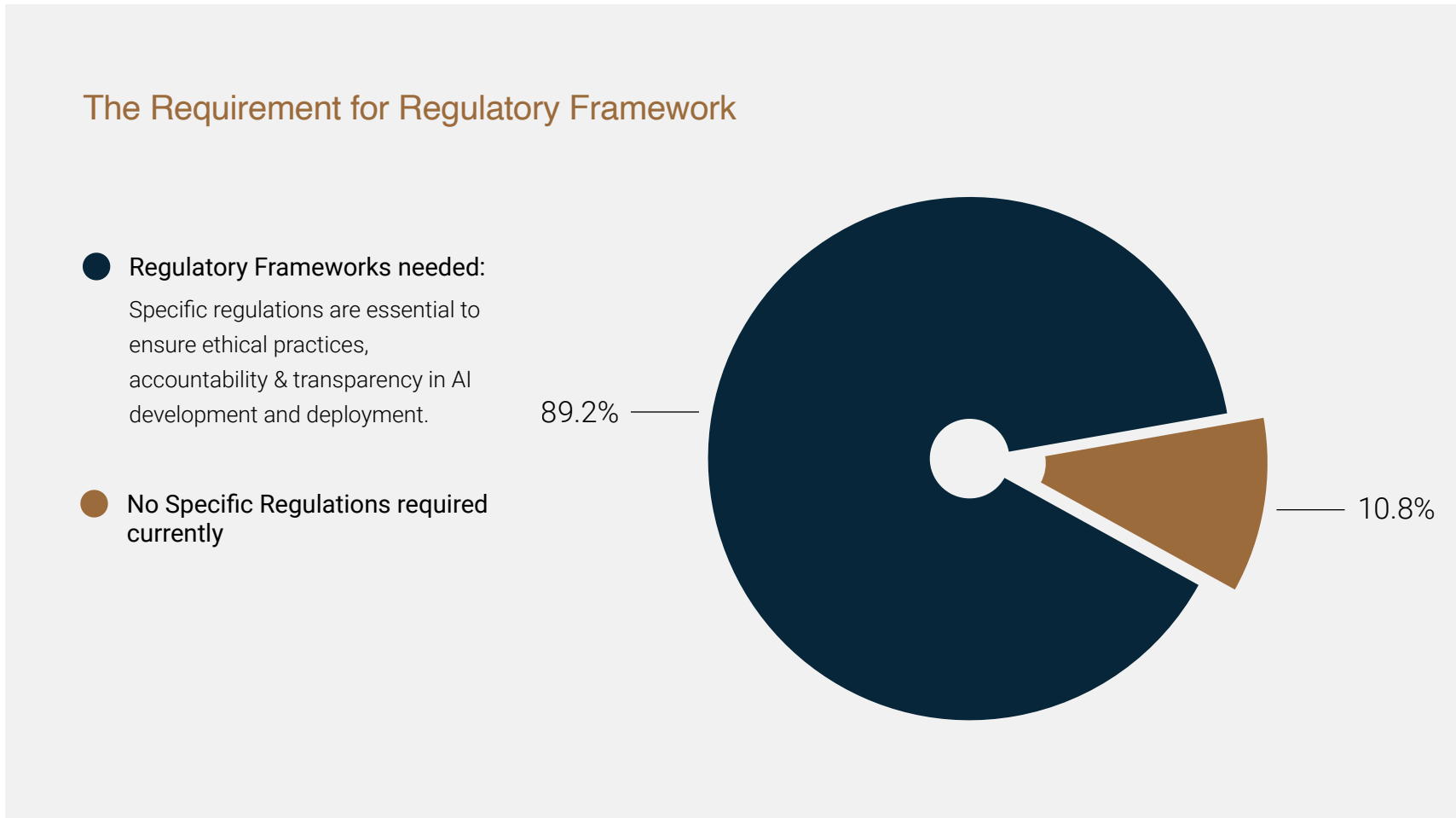
Fig. 6 | Perspective on AI Regulation in India

06

Hybrid Regulatory Approach is Emerging as a Favourable Choice over Other Routes

The industry survey indicates a strong preference, a considerable 51.4%, for a hybrid regulatory approach for AI governance in India. This approach blends traditional regulatory frameworks with adaptive quality vis-à-vis flexibility, incorporating technology-demanding moderations with time, and addressing the unique challenges posed by AI. Such a framework also balances innovation with necessary oversight, assuring AI's intended benefits while mitigating potential risks.

Such an approach will also allow for the implementation of regulatory sandboxes, enabling controlled experimentation and iterative development of AI technologies under regulatory supervision.

———

[20] https://digitalregulation.org/3004297-2/

Globally, the hybrid approach to tech and digital regulation is becoming more prevalent as it offers flexibility and adaptability. For instance, the UK's "pro-innovation" proposed framework for AI regulation is notable for its context-specific and risk-based approach. This framework, outlined in the UK Government's white paper, emphasizes proportionality and practicality, allowing different sectors to tailor regulatory measures based on the AI applications' risk level and context of use.[20] The UK's approach stands in contrast to the more prescriptive regulations seen in the EU's AI Act, highlighting the benefits of a more flexible, principle-based method .

The Preferred Approach for Regulatory Framework

A self-regulatory approach as a potential way forward — 8.1%

A separate regulatory approach with principles of ethical AI enshrined — 35.1%

A mixed/hybrid approach for regulation, tailored to and depending on the principle or type of AI — 51.4%

The current models and guidance frameworks should suffice for now — 5.4%
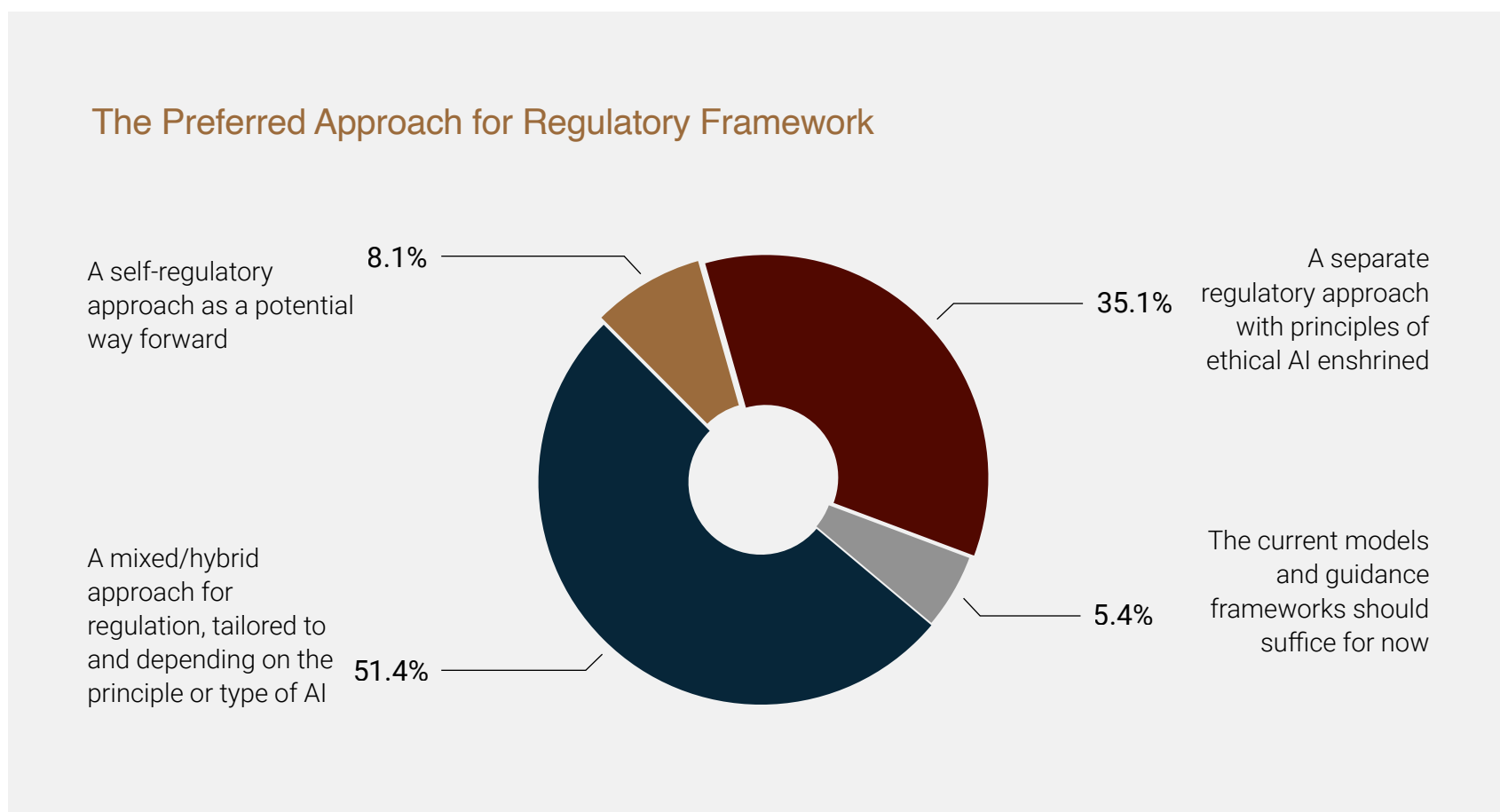
Fig. 7 | Preferred Regulatory or Governance Approach for AI in India

___

[20] https://digitalregulation.org/3004297-2/

In the absence of regulation establishing AI governance mandates in India, leaders who proactively incorporate AI governance are those that recognise the business, reputational, and social advantages in doing so. Others that kick the can along the road either find it cumbersome or miss the sustained economic and competitive advantages of proactive AI governance. Waiting for compliance to kick in could cost businesses much more than if they prioritise resources early on. And there are simpler ways to initiate AI governance than aiming to achieve high ideals from the get go.

Nidhi Sudhan

Founder - Citizen Digital Foundation | Also honoured as one of 100 Brilliant Women in AI Ethics 2024™

Navigating the Regulatory Maze: India's Responsible AI Approach

So far, no AI based regulations have been released by the government, including any official mandate on "Responsible AI". However, there have been several developments on the subject pertaining to innovating its capability, or curbing its wrongful use.

While many experts have put forward self-regulation as an option, the Economic Advisory Council to the PM however suggests[21] the presence of a **separate "agile" regulator**. Such a regulator, the EAC says, can proactively identify potential pitfalls and recalibrate regulations in real-time, ensuring that governance remains both relevant and robust.

The Digital India Bill or equivalent revisions of the Information Technology Act, 2000 (IT Act), if and when introduced, will look at AI as a separate sector and from a user harm perspective, principally in alignment with the philosophy and values of responsible and ethical use of AI. This will encapsulate three pertinent pillars[22] - **Open Internet, Online Safety and Trust, Accountability.**

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("**IT Rules, 2021**") require intermediaries, like social media platforms, to be responsible for maintaining a safe and trustworthy internet. This includes quickly removing prohibited misinformation, false information, and deep fakes. If these platforms fail to meet these requirements, they lose their safe harbour protection provided under section 79 of the IT Act.[23]

Recently enacted laws like will also have its impact on India's responsible AI ecosystem. **The Digital Personal Data Protection (DPDP)** Act will have corporations curating its data management and usage for which AI is an essential ingredient (refer Annexure II for more details). Bharatiya Nyaya Sanhita[24] (BNS) is expected to act as a propellant to combat harms, and individuals' dignity and privacy, especially in the realms of emerging technology trends like AI.

A potential draft *AI Regulation framework* is also set to be released in 2024. The intention is to harness AI for economic growth and address potential risks and harms, as quoted by MoS-MeitY. On a separate occasion, Union Minister Ashwini Vaishnaw also discussed on the possibility of AI regulation that caters to the safety and copyright issues of online news publishers and content creators. It is not clear if both the frameworks will be drafted as one, or intended to be separate.

---
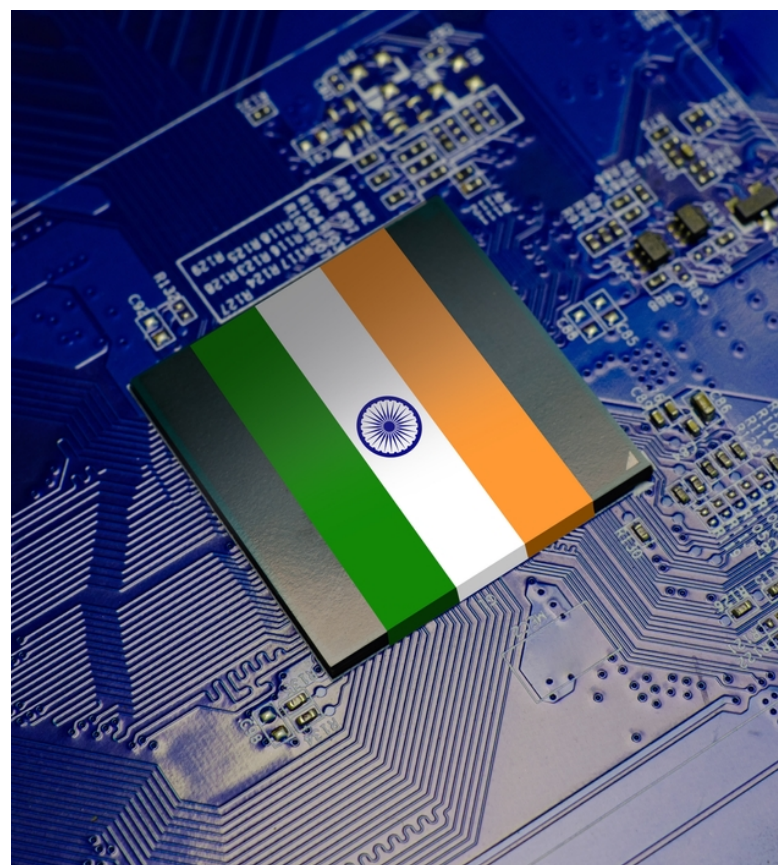
[21] https://eacpm.gov.in/wp-content/uploads/2024/01/EACPM_AI_WP-1.pdf

[22]Information collected by Primus after attending the Mumbai DIA consultation in 2023.

[23]https://pib.gov.in/PressReleasePage.aspx?PRID=2036440

[24]https://prsindia.org/files/bills_acts/bills_parliament/2023/Bharatiya_Nyaya_Sanhita,_2023.pdf

How Responsible AI is intertwined with India's Competition Law Regime?

India's current competition law framework, primarily governed by the Competition Act of 2002 and its amendment in 2023, aims to prevent practices having an adverse effect on competition, to promote and sustain competition, to protect the interests of consumers, and to ensure freedom of trade. However, this framework has not been fully adapted to address the unique challenges posed by AI, particularly in terms of algorithmic advantages, data monopolies, and the potential abuse of dominant positions.

AI technologies, by their nature, can significantly alter market dynamics. They can lead to new forms of market power, such as data monopolies, where companies with vast amounts of data can gain an unfair advantage. AI algorithms can also facilitate tacit collusion without explicit agreements between competitors, a phenomenon that traditional competition law is not equipped to handle. Moreover, AI can lead to discriminatory pricing and consumer profiling, raising concerns about consumer welfare and market fairness.

The enforcement of the DCA, the draft[1] bill of which is published, is expected to bring a more nuanced and forward-looking approach to competition around digital enterprises in India. An ex-ante regulatory framework, as proposed in the DCA, would allow for proactive monitoring and regulation of AI-driven market practices.

This includes regulations on data accumulation and use, amongst others, ensuring that AI-driven businesses do not abuse their market position. The rules emanating from the Act might also involve setting or altering standards for transparency and fairness in AI algorithms to prevent anti-competitive practices such as price-fixing or market manipulation. CCI announcement for conducting a market study on the impact of AI on businesses and services, or the bill including provisions for research with external agencies, are steps in that direction. However, the proposed ex-ante framework is facing criticism from the industry with the concern that it will stifle innovation, and put barriers to products and services that consumers find otherwise convenient to use.

Our Recommendations: Principle driven Frameworks for Responsible AI in India

01 **Approach based on** Risk Category

This approach outlines the evaluation and governance of AI systems based on the principles of Responsible AI, considering different risk categories and types of AI. The following matrix provides a nuanced view of **how different AI systems might be evaluated and governed against the identified principles of Responsible AI**, considering a wider range of risk categories and types of AI. Each principle is checked against the type of AI system where it is most relevant, but this does not mean that unchecked principles are not applicable to those systems; rather, **the checked principles are those with a significant relevance or impact**. Also, and as an example, for medium and low risks entries, the approach can be via adequate Code of Conduct, while for High-Risk entries, a more robust (regulatory) oversight similar to EU's approach is an option. The matrix has been formulated after careful consideration and extensive discussion with industry experts, thought leaders, and AI innovators.

No.	Principles of Responsible AI	High Risk Use Cases (e.g., Health-care Systems, (Mis)-information)	Medium-Risk Use Cases (e.g., LLMs, Online Ads)	Low-Risk Use Cases (e.g., Chatbots, Search Results, Video Games)	Predictive Analytics	Autonomous Systems (e.g., Self-driving Cars, Smart Factories)	Social Media Algorithms
1	Inclusive Growth, Sustainable Development and Well-being	✓			✓	✓	
2	Human-centered Values and Fairness	✓	✓		✓		✓
3	Transparency and Explainability	✓	✓	✓	✓		✓
4	Robustness, Security and Safety	✓	✓		✓	✓	
5	Accountability	✓	✓	✓	✓	✓	✓
6	Privacy and Data Governance	✓	✓		✓	✓	✓
7	Innovation and Scalability	✓	✓		✓	✓	
8	Response and Redress Mechanisms, including a Risk Management System	✓	✓	✓	✓	✓	✓
9	Human Oversight	✓	✓			✓	
10	Intellectual Property	✓	✓	✓	✓	✓	✓

Table 1 | Approach based on Risk Category

02   Approaching according to the Sector

India is already making sectoral progress in **Finance** (Securities and Exchange Board of India's (SEBI) directives[25]) and **Healthcare** (Indian Council of Medical Research (ICMR) published the "Ethical Guidelines for AI 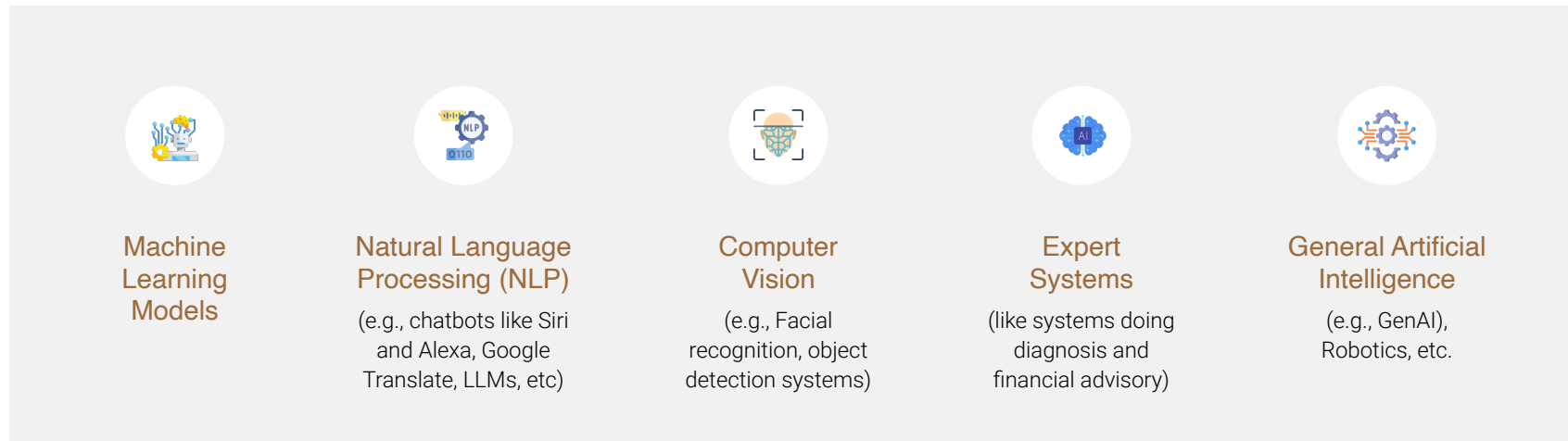in Biomedical Research and Healthcare") at incorporating values of AI's ethical practice and implementation. More sectors need requisite intervention, such as **Industrials & Automotive, Retail, FMCG/CPG, Fintech and Agri-tech**, which are expected to drive more than 60% of AI's Gross Value Added contribution in India's GDP by 2025.[26] Thus, sectoral specific guidelines or framework can be another approach.

Industrials & Automotive	Retail & FMCG/CPG
Fin-Tech	Agri-Tech
Finance	Healthcare
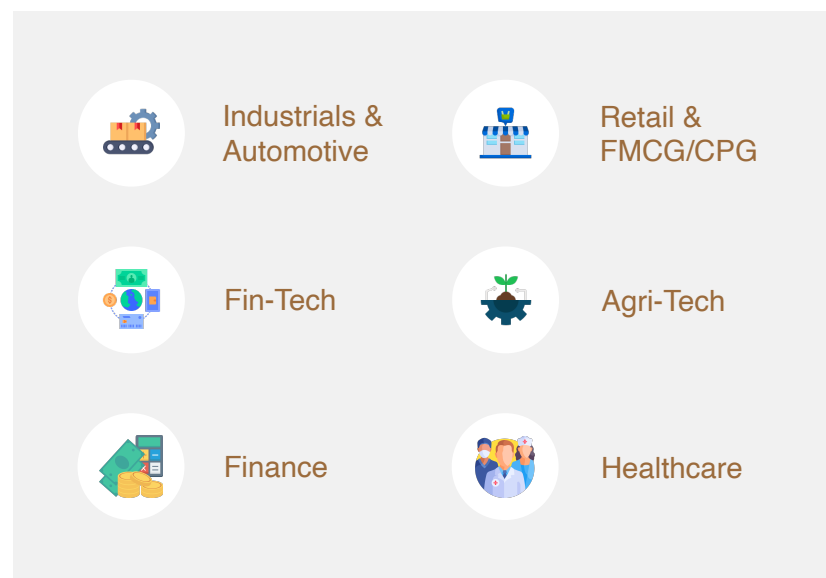
03   Approaching according to the Type or Underlying technology or Foundational models

While India still awaits its national level policy or guidance on AI use (like EU's AI Act), the responsible application of the technology and its principles can be voluntary which businesses offering different type of AI models should strive to adopt. Approach can be thus different basis the type of **AI model or the underlying technology used**; that is, not strictly on the risk factor but on the digital infrastructure applied –

Machine Learning Models	Natural Language Processing (NLP)	Computer Vision	Expert Systems	General Artificial Intelligence
	(e.g., chatbots like Siri and Alexa, Google Translate, LLMs, etc)	(e.g., Facial recognition, object detection systems)	(like systems doing diagnosis and financial advisory)	(e.g., GenAI), Robotics, etc.

The separate framework of each type could ensure a more focused development of the technology. It may however be a disadvantage with the scope of AI now getting perpetually infinite, and with more innovations and inventions, will lead to a requirement of more frameworks.

Another school of thought is that while regulating foundational models, although a considerable proposition, can be a tricky scenario. One of the key considerations here is the balance between promoting innovation and mitigating systemic risks while providing clear regulatory requirements. This balance is crucial because foundational AI models, unlike their predecessors, can be applied across a broad variety of use cases, that is, it can be the foundation for many applications of other AI models. This broad applicability raises new challenges in terms of regulation, as these models have the potential to fundamentally change many areas of life and introduce substantial systemic disruptions. As policymakers begin to regulate AI, it will become increasingly necessary to distinguish clearly between types of models and their capabilities, and to recognise the unique features of foundation models that may require additional regulatory attention.

———

[25] https://www.sebi.gov.in/legal/circulars/jan-2019/reporting-for-artificial-intelligence-ai-and-machine-learning-ml-applications-and-systems-offered-and-used-by-market-intermediaries_41546.html and https://www.sebi.gov.in/legal/circulars/may-2019/reporting-for-artificial-intelligence-ai-and-machine-learning-ml-applications-and-systems-offered-and-used-by-mutual-funds_42932.html

[26] https://www.investindia.gov.in/team-india-blogs/artificial-intelligence-powering-indias-growth-story#:~:text=60%25%20of%20AI's%20Gross%20Value,important%20application%20areas%20for%20AI.

Efficacy of Governing Open Foundational Models

The availability of foundation models varies: some are fully closed (e.g., Google DeepMind's Flamingo), some offer hosted access (e.g., Inflection's Pi), others provide cloud-based access via API (e.g., OpenAI's GPT-4), cloud-based fine-tuning (e.g., OpenAI's GPT-3.5), widely available weights (e.g., Stability AI's Stable Diffusion, Meta's Llama 2), or complete access with weights, code, and data (e.g., BigScience's BLOOM with restrictions, EleutherAI's GPT-NeoX without restrictions).

Open foundation models benefit society by fostering competition, accelerating innovation, and decentralizing power. Smaller companies can choose from various open-source models, which are typically more affordable and offer increased transparency and accountability. Releasing models with training data allows independent third parties to assess capabilities and risks better.

However, open foundation models pose potential risks. Open developers cannot effectively restrict malicious actors from removing safety guardrails. Restrictions on downstream use are often ignored, increasing the risk of generating disinformation, cyberweapons, and spear-phishing emails. Legislation making developers liable for content produced by their models could harm open-source projects, as users might alter models to create illegal content. Policymakers should carefully consult with open foundation model developers to avoid unintended consequences and consider whether placing liability for downstream use is an effective governance approach.

Indian government (MeitY) may also develop its own foundational model which will be customised for use by Indian companies, entrepreneurs, academics and researchers. The initiative is to be led by the IndiaAI Innovation Centre to be set up by MeitY under the previously stated IndiaAI Mission. It could be a large action model (LAM) or large multimodal model (LMM) so that the output can be used for a wide range of applications and services.

Every single walk of life will be impacted by AI-based algorithms, and this will be far beyond the impact that machines or computers had on the world in the past. AI can, if regulated properly, be an obedient servant. Regulation must be done by private bodies which frame ethical codes which are coded into, and driven by, the AI-based algorithms themselves. Regulatory bodies must be transparent, responsive and nimble. Heavy-handed and strict regulations which are sought to be implemented in a top-down manner will be ineffective.

Srinath Sridevan
Senior Advocate, Madras High Court and Founding Partner, HSB Partners

Role of Government Enforcement Mechanisms

This following table presents a foundational view and could be expanded with deliberations for more specific roles and responsibilities as the regulatory landscape for AI in India develops. The suggested RAI Accountability Committee, in particular, is a suggested body that could play a significant role in bridging gaps between different regulatory entities and ensuring a comprehensive approach to Responsible AI governance.

Table 2 | Role of Govt. Enforcement Mechanisms

🟦 Bodies already present or proposed to be established soon

⬜ Bodies or enforcement mechanisms suggested

Enforcer/ Oversight Body	Details	Function / Responsibility
Data Protection Board (DPB)	The to be established DPB under the DPDP Act	• Ensure AI compliance with data privacy and protection laws. • Manage data breaches and ensure transparency in AI operations.
CCPA (Dept. of Consumer Affairs)	CCPA to introspect and direct orders from a consumer standpoint	• Protect consumer rights in AI applications. • Address AI-related product/service issues. • Ensure non-discriminatory AI practices.
Competition Commission of India (CCI)	CCI can look at AI from competition perspective, with the upcoming Digital Competition Act already in the pipeline	• Monitor and regulate AI-driven market practices to prevent anti-competitive behaviours. • Ensure that AI technologies do not create monopolistic or unfair market conditions. • Investigate and address issues related to AI-driven price fixing, market manipulation, or exclusionary practices. (algorithmic collusion) • Managing and safeguarding IPR.
RAI Accountability Committee/ National AI Ethics Committee	The committee to be formed as an inter-ministerial one, or under a relevant ministry working in coordination with other departments	• Regularly review AI ethics and compliance. • Publicly report AI impacts. • Coordinate among regulatory bodies for AI governance. • Develop national AI ethics like guidelines, based on risks/ sectors/ type of AI. • Monitor AI research and development for ethical adherence. • Facilitate collaboration across ministries for AI policy implementation. This also includes assessing relevance of responsible AI use across different sectors and industries. • Ensure that, with regard to high-risk AI systems especially, there is appropriate coordination and cooperation. • Establish and timely update a database of different risk systems. • Establish, after prior approval of the Central government, scientific sub-committees, to oversee specific technical matters. • Ensure unified approach to AI governance. • Oversee PPP models/ industry collaborations on AI. • - Ensure AI models and responsible guidelines are proportionately adopted and adhered by organisations.[27]
AI Incident Response Team	A suggested separate and dedicated team under the aegis of CERT-In	• Quickly respond to AI-related incidents or breaches. • Implement corrective actions and updates in AI systems.
AI Technology Audit & Certification Body	A separate body under BIS is suggested	• Certify AI systems or suggest standards for ethical compliance. • Conduct regular audits of AI technologies in use, in coordination with other authorities.

[27] https://www.wired.com/story/most-news-sites-block-ai-bots-right-wing-media-welcomes-them/?utm_source=bensbites&utm_medium=newsletter&utm_campaign=daily-digest-next-gen-ai-videos

07

Sectoral Study: Responsible AI in Online Advertising

It has been well established that online advertisements garner a significant amount of revenue for tech giants and companies[28]. In view of that, it is important to dive into the applicability of responsible AI in the sector. Powered by complex algorithms, personalised ads anticipate one's desires, predict purchases, and weave themselves seamlessly into her/his digital experience. But within this algorithmic web, lies the potential for bias, discrimination, and manipulation. This is where responsible AI steps in, not as a constraint, but as a guiding innovation towards ethical and impactful online advertising.

Consider *Rani*, a young aspiring musician browsing music equipment websites. Unbeknownst to her, the AI powering the ad platform has identified her gender as female. Consequently, she is bombarded with ads for beginner instruments and pink-hued accessories, while her male counterparts encounter promotions for professional-grade equipment and technical tutorials. This algorithmic bias, though unintentional, perpetuates gender stereotypes and limits Alice's access to information and opportunity.

Again, imagine *Mohan*, a recent immigrant struggling to navigate the financial complexities of his new home. The AI, drawing on incomplete data and potentially biased credit scoring models, excludes him from targeted ads for affordable loans and financial services. This algorithmic discrimination exacerbates his economic vulnerability and widens the digital divide.

These scenarios highlight the crucial role of responsible AI in online advertising. To counter bias, algorithmic fairness frameworks can be implemented. These frameworks involve auditing training datasets for discriminatory patterns, employing counterfactual reasoning to assess the impact of algorithmic decisions on different demographics, and implementing techniques like adversarial training to mitigate bias in model outputs.

Moreover, explainability and transparency become paramount. Users deserve to understand why they see specific ads and have the ability to contest decisions based on algorithmic bias. Explainable AI (XAI) techniques can be employed to demystify the black box of algorithms, revealing the factors influencing ad targeting and allowing users to challenge unfair inferences.

Privacy concerns also come to the fore. To build trust and foster user autonomy, privacy-preserving ad targeting methods become essential. Federated learning, where models are trained on decentralized datasets without compromising individual data, and differential privacy, which adds controlled noise to data to protect sensitive information, offer promising solutions. Finally, user control reigns supreme. Granular settings allowing users to opt-out of personalized ads, curate their ad preferences, and manage data sharing empower them to navigate the online landscape with agency and informed consent.

_____

[28] https://www.statista.com/statistics/205352/digital-advertising-revenue-of-leading-online-companies/
#:~:text=In%202021%2C%20Facebook%20earned%20over,dollars%20through%20digital%20advertising%20channels.

While these principles may seem abstract, Google's "Why this ad?" feature and ad control settings offer concrete examples of implementing responsible AI in online advertising. Moreover, Google's larger Responsible AI principles[29] are effectively implemented in all its business practices. However, the journey extends far beyond one platform. Industry-wide collaboration, standardised ethical frameworks, and continuous research are critical to ensuring responsible AI becomes the bedrock of online advertising, fostering a digital ecosystem that is fair, transparent, and empowering for all.

Over the next decade, smart assistants/tools like Alexa or a Siri will transform how companies sell to and satisfy consumers, and global firms will battle to establish the preferred artificial intelligence platform. These AI tools will become trusted advisers to consumers, anticipating and streamlining their purchase needs. That is, there could be a scenario where consumers will align their shopping trends and preferences not based on brands but what the AI tools suggest them with.[30]

Additional Point for Further Deliberation

Large technological companies, whose primary revenue comes from providing platforms for online advertisements, act solely as technological intermediaries and do not determine the specific content of advertisements. For example, while a tech platform may support inclusivity in a beauty product ad featuring a woman, the final content is created by the advertisers. The intermediary is not responsible for societal perceptions, such as the notion that fair-skinned women are more beautiful. Therefore, there is a need to address and improve the 'humanness' aspect in advertising.

[29] https://ai.google/responsibility/principles/
[30] HBR's 10 must reads on AI, Analytics, and the new Machine Age. (2019). Harvard Business Review

Case Example

Google's Inclusive AI Model

Google is taking a credible step towards fairer tech with their Monk Skin Tone Scale (MST Scale), being developed by advanced AI and ML. This 10-shade system, co-created with sociologist Dr. Ellis Monk, aims to break free from limited representations and embrace the beautiful spectrum of human complexions.

The MST Scale is being woven into Google's fabric, from refining image searches to evaluating facial recognition tech to making online ads more credible. This means no more one-size-fits-all algorithms, but nuanced understanding and representation for all. They're also working on inclusive labelling systems and fostering open collaboration to keep the MST Scale evolving.

This scale has the potential applicability in significantly enhancing online advertising by offering a nuanced representation of diverse skin tones. This 10-shade system allows for more accurate and inclusive targeting, ensuring that a wide range of ethnicities and races are represented in advertisements. This improved representation leads to more personalized and relevant ads, particularly in sectors like beauty and skincare, where skin tone is a critical factor. By incorporating the MST Scale, biases inherent in traditional AI algorithms can be reduced, promoting fairer ad distribution and avoiding the pitfalls of one-size-fits-all approaches. Additionally, this approach demonstrates a commitment to ethical standards in AI, addressing cultural sensitivity and compliance issues.

In Dr. Monk's words, "In our (Google's) research, we found that a lot of the time people feel they're lumped into racial categories, but there's all this heterogeneity with ethnic and racial categories…we need to fine-tune the way we measure things, so people feel represented."

Image and information sourced from https://blog.google/products/search/monk-skin-tone-scale/

Conclusion and Way Forward

Review of Key Listed Recommendations

Data and Technological Infrastructure

01   ## Data Management

In India, prioritizing data quality and availability in key sectors will greatly enhance AI applications. Establishing strong data governance frameworks and investing in advanced infrastructure will ensure data integrity and efficient data handling. Promoting interoperable standards will enable seamless data sharing, while innovative collection techniques like IoT and remote sensing will improve data quality. Additionally, investing in training programs for data management skills will further boost data availability and quality.

02   ## Computational Resources

India's AI ecosystem is growing, but the availability of computational resources remains a bottleneck. Investments in infrastructure, including setting up AI research centers and supercomputing facilities, are essential to address these challenges.

03   ## Explainability and Transparency

Institutionalising transparency and accountability in AI systems for fostering trust and accountability. This can involve mandating 'data cards' or model cards (similar to nutrition levels in food). Transparency can also be on the governance structures of companies, and around data sharing. Discussion on disclosure of how algorithms work, criteria used in decision-making, and the implications for end-users should also be taken up as a priority, to counter unintentional bias and impact on digital creators.

04   ## Curbing unaccounted and misleading deepfakes

Strong measures are required to protect personal data and maintain the integrity of AI-generated content. Established tech platforms can help government integrate deepfake monitoring with their current mechanisms of addressing misinformation. As an example, initiatives like 'PIB Fact Check' or 'DigiKavach' can encompass a deepfake alert infrastructure. Geopolitically, deepfakes also need to be looked at from a cross-country liability chain intervention instead of national silos.

05   ## Emerging tech like blockchain

Emerging and future technologies hold immense possibilities, and can enhance the transparency, security, management and traceability of AI systems, thereby supporting more robust and trustworthy AI applications.

Policy and Regulatory

06   A dedicated and focused AI ethics board within the corporate governance structure

This involves the leadership who continually work to ensure that AI systems are functioning properly, safely, and responsibly. When an AI system for credit approval, for example, is found to be discriminating against certain groups, ethics managers or the 'Ethics Board' can be responsible for investigating and resolving the issue.

07   A Robust Yet Hybrid Regulatory Approach

This approach blends traditional regulatory frameworks with adaptive quality vis-à-vis flexibility, incorporating technology-demanding moderations with time, and addressing the unique challenges posed by AI. Such route also balances innovation with necessary oversight, assuring AI's intended benefits while mitigating potential risks. This can be further informed by the emerging global best practices. On the framework, it can be based on Risk Category; or the Sector the AI tool is serving; or Type or Underlying technology or Foundational models.

08   Enforcement mechanisms suggested other than already proposed by the government

An RAI Accountability Committee/ National AI Ethics Committee, A dedicated AI Incident Response team under the aegis of CERT-In, and an AI Technology Audit and Certification Body, are recommended, other than bodies like the Data Protection Board, the CCI and the Central Consumer Protection Authority, all carrying their respective functions and duties.

09   Institutional capacity

To truly leverage institutions focusing on AI, there needs to be a concerted effort towards cohesive ethical standards for AI, aggressive private investments, skilling and continuous reskilling of the talent pool, and a sustainable roadmap for AI development in India.

Responsible Adoption

10 Addressing bias in language use

Discussions are required on the 'subjectivity' of what bias encompasses. This becomes especially important on building the specifications of biases around India's diverse language use. Handholding startups who are working on different language models will be an essential criterion.

11 Voluntary Adoption - Secure by Design

Incorporating a 'secure by design' approach in AI systems is critical. This involves integrating advanced security features at every development stage to protect against threats like data breaches and cyberattacks. By ensuring robust encryption, authentication, and continuous monitoring, AI systems can enhance trust and reliability, safeguard sensitive data, and promote wider adoption of AI innovations.

12 Voluntary Adoption – Responsible by Design

For AI developers and deployers, voluntarily adopting a 'responsible by design' approach to ensure ethical standards and practice are integrated from the outset, enhancing the trust and safety of AI applications. This proactive commitment demonstrates leadership and responsibility in aligning technological advancements with societal well-being.

Integration and Collaboration

13 Integration with Existing Systems and Processes

Successful integration of AI into business operations can drive efficiency and innovation. In India, sectors like finance, healthcare, and manufacturing are already seeing AI integration, but a broader institutional level vis-a-vis more coordinated effort is needed to overcome technical and organizational barriers.

14 Cross-Border (Digital) Trade

For India to become a global AI hub, it must navigate such restrictions effectively. Strengthening international partnerships and advocating for open data exchange can help mitigate these barriers and promote a collaborative AI research environment.

15   Addressing the skill gap

India's AI skill gap highlights the necessity for collaborative efforts between the government, academia, and the corporate sector to design comprehensive skill development programs and promote continuous education in emerging technologies like AI. The adequate and proportional channelisation of young and skilled tech talent in STEM towards AI should also be a high priority.

16   R&D

Prioritizing research and development in areas like bias detection, explainability, and privacy preserving techniques will equip India with the necessary technical know-how to mitigate potential risks. This is also crucial for continuing the momentum of India's patent filing track in AI.

17   Consumer Awareness

A dedicated 'AI Consumer and Education Program', as a joint inter-departmental activity can be a start. This could include integrating AI education into school and university curricula, and initiating community engagement programs for interactive learning via CSCs.

In conclusion, to fully unlock the transformative potential of AI while safeguarding ethical considerations, **fostering collaboration** among stakeholders is paramount. Stakeholders include governments, intergovernmental organizations, the technical or developer community, civil society, researchers, academia, media, education policymakers, private sector companies, human rights institutions, equality bodies, anti-discrimination monitoring bodies, and groups for youth and children.[31]

---

[31] https://unesdoc.unesco.org/ark:/48223/pf0000381137
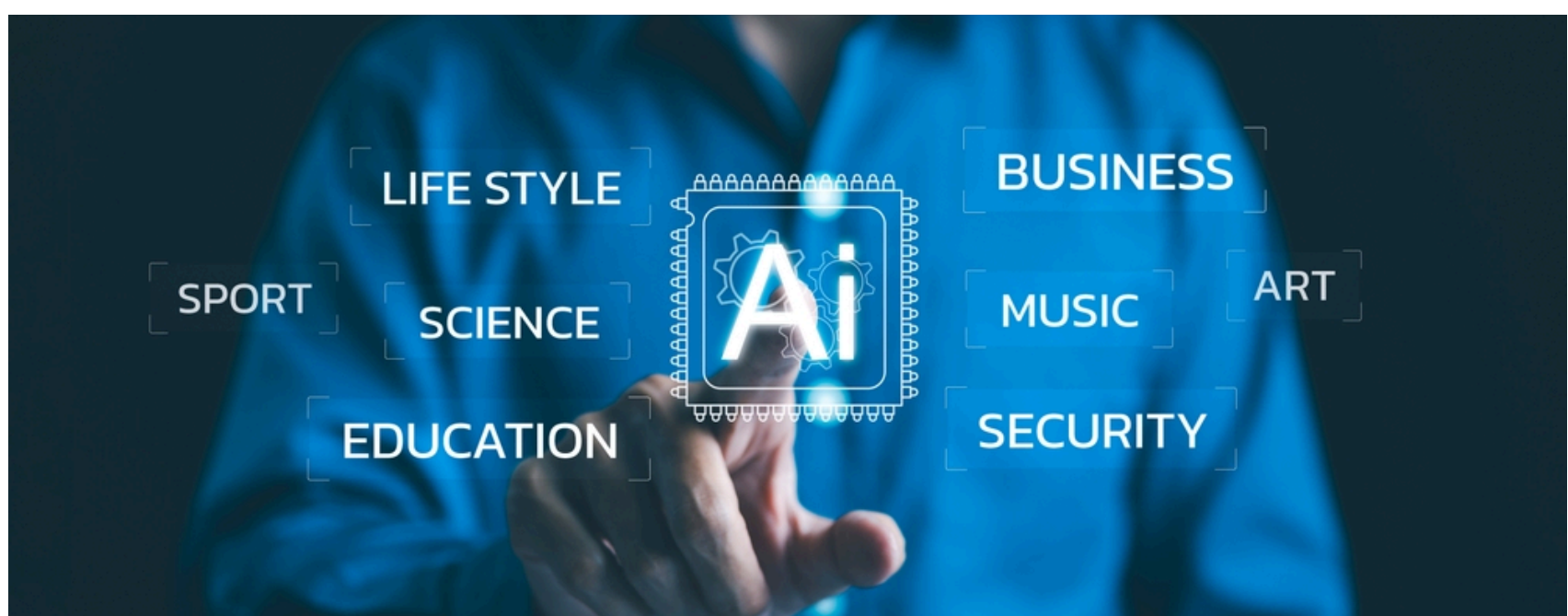
The Advantage of Collaborative Intelligence

What is instinctive for humans, like telling a joke, can be challenging for machines, while tasks that are simple for machines, such as processing vast amounts of data, are nearly unfeasible for humans (also referred to as Moravee's Paradox). Both levels of abilities are essential in business.

Many companies have utilized AI to automate tasks, yet those that primarily use it to replace employees will experience only short-term increases in productivity. Furthermore, there have been no studies to prove that there are long-term gains after such an action.

Companies rather gain from enhancing the cooperation between humans and artificial intelligence. There are five principles to facilitate this: rethinking business processes; promoting experimentation and employee participation; guiding AI strategy consciously; collecting data ethically; and restructuring work to integrate AI while developing the necessary employee skills. A survey across 12 industries revealed that the more of these principles companies implemented, the more successful their AI initiatives were in terms of speed, cost efficiency, revenue, and other operational metrics.[32]
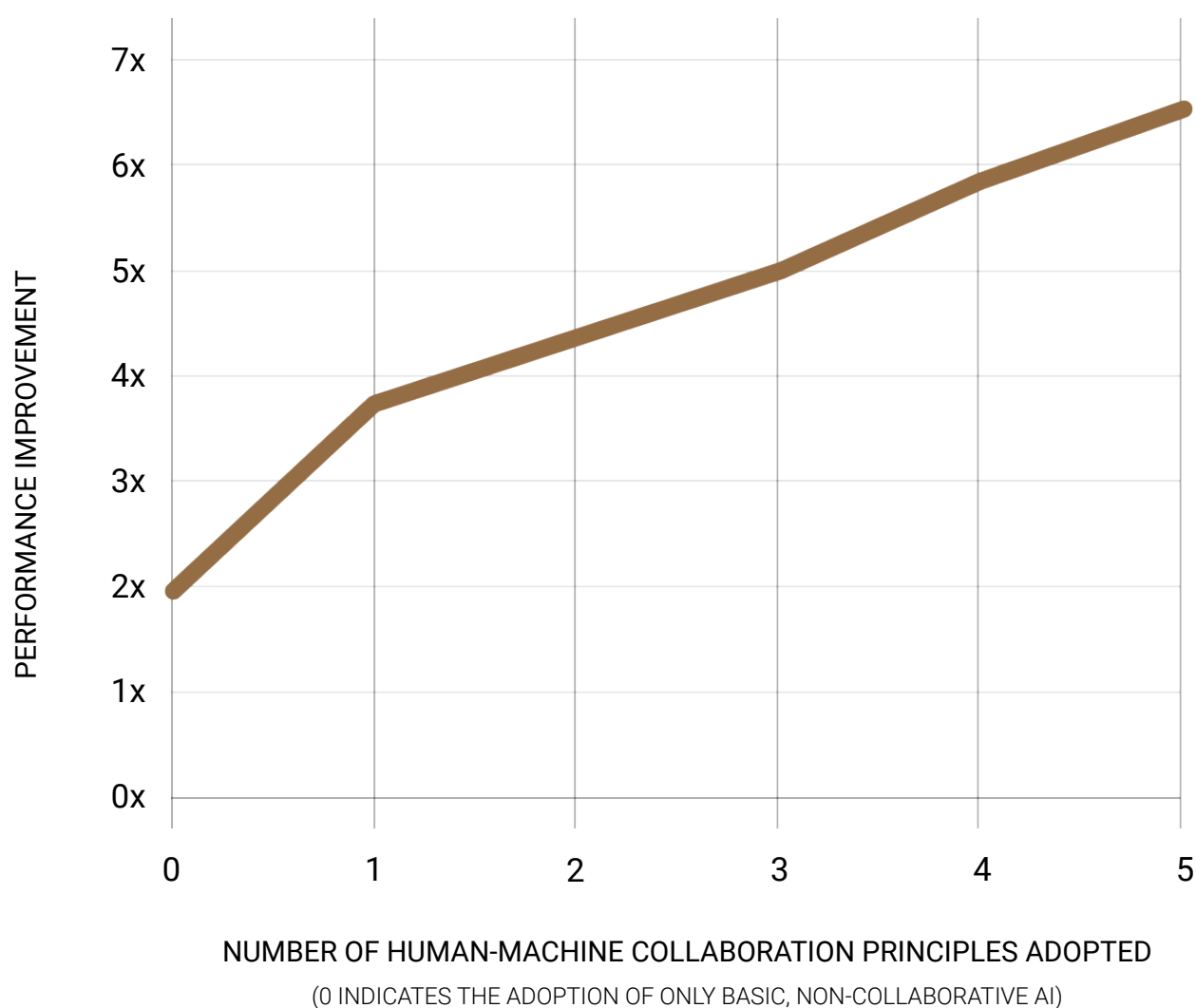
Fig. 8 | The Advantage of Collaborative Intelligence

Companies need to grasp how humans can best complement machines, how machines can improve human capabilities, and how to reshape business processes to foster this cooperation. Research and practical experience have led us to consider the importance of this 'collaborative intelligence'.

———

[32] https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces

Annexure I
Research
Methodology

Mixed Methodology Approach

The study adopts a mixed methodology approach, which integrates both qualitative and quantitative research methods. This approach allows for a holistic exploration of the subject matter, providing a more comprehensive view of the experiences and perceptions of the industry regarding responsible AI.

Questionnaire-Based Sample Analysis

The foundation of this study is the utilization of a structured questionnaire. The survey questionnaire was meticulously designed and prepared by Primus, after consulting subject matter experts. It was implemented vis-à-vis disseminated by Primus. This questionnaire served as the primary tool for collecting data from the target audience.

Sampling and Stratification

The study employed a rigorous sampling technique to ensure the representativeness of the sample population. Relevant stratification criteria were identified, leading to the division of the sample population into homogeneous groups or strata. These strata included geographical divisions, industry types and industry stage (established, startup), among others. This approach allows for the examination of variations and patterns within specific segments of the technology sector.
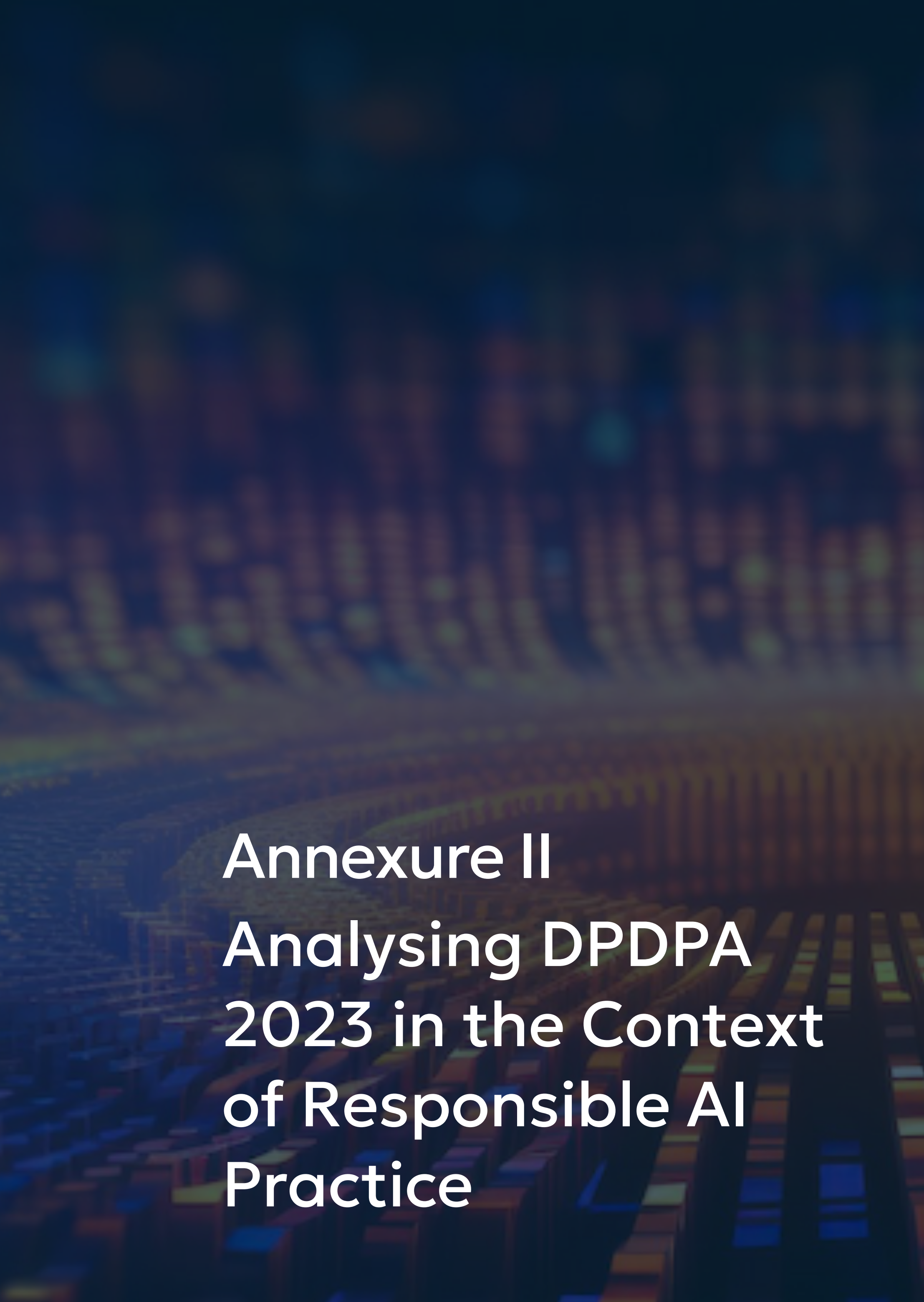
Data Collection

Data collection involved the systemic administration of the questionnaire to the selected industry leaders, wherein 75 responses were registered. The data collection process was conducted primarily through online surveys.

Analysis of Perceived Impact

To determine how the industry perceives the impact and rise of responsible AI, the survey inquired about various aspects including decision making, challenges around data and privacy, key principles, human resource, operational efficiencies, scope of regulation, etc. This information was collected through the questionnaire responses. Subsequently, Primus conducted rigorous statistical analyses to quantify these aspects.

Annexure II

Analysing DPDPA 2023 in the Context of Responsible AI Practice

Provisions emanating from the Act	Impact on Businesses using AI & Interpretation for Responsible AI	Primus Point-of-View
Consent for Data Processing	AI businesses must obtain explicit and informed consent for data processing. This impacts how AI algorithms are designed to collect and use data.	A shift to more user-centric design may be required.
Data Minimization	Businesses must ensure that AI systems collect only data that is necessary for their purpose, which may limit the scope of data analysis and further product development.	This could challenge AI models that rely on large datasets. The provision will also further imply curbs in 'profiling' or user behaviour pattern data to the extent of identifying an individual.
Data Localization	If the Act requires data localization, AI businesses may need to adjust their data storage solutions to comply with geographical restrictions (in reference to a 'negative list' of territories that are to be published).	This could increase operational costs for businesses, while also provide benefits of cheap labour and reduced operational costs.
Transparency in AI Operations	Businesses might have to disclose AI decision-making processes, which could affect proprietary algorithms and intellectual property.	Balancing transparency with protection of trade secrets will be key.
Accountability for AI Decisions	Businesses may be held accountable for decisions made by AI, which requires robust auditing and oversight mechanisms.	Ensuring accountability could lead to increased governance structures.
Data Subject Rights	AI businesses need to respect the rights of individuals, including access to data and the right to be forgotten, which impacts data retention policies.	Procedures for data access and deletion will need to be streamlined.
Data Protection Impact Assessments	Regular assessments may be mandatory, impacting the initial design and ongoing operation of AI systems.	May encourage more responsible AI practices from the outset.
Breach Notification	In case of data breaches, businesses must notify authorities and affected individuals, which requires strong security and incident response plans.	Cybersecurity measures will need to be robust and responsive.
Non-Discrimination	AI systems must be designed to avoid discriminatory outcomes, impacting algorithm development and training data selection.	Bias detection and mitigation will become crucial.

Table 3 | Analysis of Provisions of DPDP 2023 in context of Responsible AI Practice

Annexure III
Global AI Outlook

USA

Various federal laws, bills, executive orders and frameworks on AI have been passed or proposed on a range of matters (from strategic or administrative initiatives to regulation of specific AI issues or applications such as deepfakes and foundation models).

A key landmark development was Executive Order 14110 (Nov 2023) which established a comprehensive suite of regulatory, economic and strategic initiatives in AI. Some US states have direct AI regulation, with the top 3 states (in terms of number of AI-related bills passed in 2023) being California, Virginia and Maryland. U.S. AI Safety Institute Consortium (AISIC) has also been proposed, which aims to unite AI creators, users, academics, government and industry researchers, and civil society organizations to support the development and deployment of safe and trustworthy artificial intelligence.

UK

Regulation by sector approach. The UK relies on existing regulators to interpret, implement and enforce various AI ethics principles (set out in the National Whitepaper on AI Regulation) in their relevant sector. It hosted the world's first Global AI Safety Summit (2023), establishing the Bletchley Declaration and UK AI Safety Institute. The country intends to take primarily a "pro-innovation" stance.

Brazil

Brazil's Bill proposes rights for the people affected by AI systems, proposes different levels of risk for AI system, mandates transparency and bias mitigation, and Algorithmic Impact Assessments (AIA) before deploying high-risk AI systems.

● Countries with AI Regulation, Policy or Frameworks, either implemented or in discussion

● Countries without AI Regulation or any credible Policy Framework

Fig. 9 | Global AI Outlook

European Union (EU)

The recently approved EU AI Act is the first legally governed AI legislation passed in the world. This act mandates greater transparency from companies regarding data usage and imposes new restrictions on AI use cases, such as banning the creation of facial recognition databases and using emotion recognition technology in certain contexts. The EU is also working on the AI Liability Directive, which aims to provide financial compensation to individuals harmed by AI technology.

China

China has launched a 3-year action plan that targets AI standards. AI regulation here is more fragmented, with specific rules for different AI applications like algorithmic recommendation services, deepfakes, and generative AI. However, in January 2024 industry ministry issued draft guidelines for standardising the AI industry.

Singapore

Singapore has published the finalised version of its Model AI Governance Framework for Generative AI. It also has an Advisory Council on the Ethical Use of AI and Data, established to advise the government on developing and implementing ethical AI governance.

Australia

The government announced that it will set up an advisory body (to be now revamped as a permanent body) of industry and academic experts to work with government to devise a legislative framework around 'high risk' AI applications. The country is reportedly considering two visions of regulation for high-risk AI: a prescriptive EU-style AI law or legislation relying on broad principles.

Acknowledgements

Primus Partners thanks the experts and organisations who were a part of the discussions, and who provided their thoughtful quotes. It also expresses its deepest gratitude to the esteemed organisations and professionals who were a part of the reviewing process of the report.

Authors & Key Contributors

Devroop Dhar
Managing Director
Primus Partners

Sayantani Chatterjee
Vice President
Primus Partners

Indranuj Pathak
Manager - Public Policy
Primus Partners

Samyak Jain
Consultant
Primus Partners

Disclaimer

The report is prepared using information of a general nature and is not intended to address the circumstances of any particular individual or entity. The report has been prepared from various public sources and the information receives from these sources is believed to be reliable.

The information available in the report is selective and subject to updation, revision and amendment. While the information provided herein is believed to be accurate and reliable, Primus Partners Private Limited does not make any representations or warranties, expressed or implied, as to the accuracy or completeness of such information and data available in the public domain.

While due care has been taken while preparing the report, Primus Partners Private Limited does not accept any liability whatsoever, for any direct of consequential loss arising from this document or its contents.

We do not claim ownership over the images used in this document.

PRIMUS PARTNERS®

Primus Partners has been set up to partner with clients in 'navigating' India, by experts with decades of experience in doing so for large global firms. Set up on the principle of 'Idea Realization', it brings to bear 'experience in action'. 'Idea Realization'—a unique approach to examine futuristic ideas required for the growth of an organization or a sector or geography, from the perspective of assured on ground implementability. Our core strength comes from our founding partners, who are goal-oriented, with extensive hands-on experience and subject-matter expertise, which is well recognized in the industry. Our core founders form a diverse cohort of leaders from both genders with experience across industries (Public Sector, Healthcare, Transport, Education, etc.), and with varied specialization (engineers, lawyers, tax professionals, management, etc.).

Navigating India

BENGALURU
91 Springboard
Business Hub 175, 176
Bannerghatta Rd,
Dollars Colony,
Bengaluru – 560076

CHANDIGARH
2nd Floor, Netsmartz,
Plot No. 10, Rajiv
Gandhi Chandigarh
Technology Park,
Chandigarh – 160019

CHENNAI
147, Pathari Rd, Door #3,
WorkEz Hansa Building,
RK Swamy Centre,
Thousand Lights,
Chennai, TN - 600006

DELHI
1 to 7, UG Floor,
Tolstoy House,
Tolstoy Road,
Connaught Place
New Delhi - 110001

KOLKATA
Siddhartha Apartments
4th Floor, 188/2,
Block J,
New Alipore,
Kolkata - 700053

MUMBAI
601, 6th floor,
Raheja Centre,
Nariman Point,
Mumbai,
MH - 400021

www.primuspartners.in   Primus Partners India   @partners_primus   info@primuspartners.in