

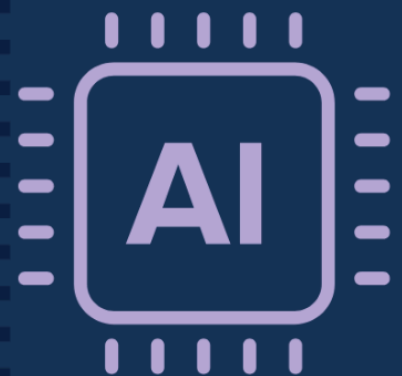


EVENT READBACK

# Shaping AI Futures: Blueprints for the Responsible AI Ecosystem in India

---

30th September 2024





# Shaping AI Futures: Blueprints for the Responsible AI Ecosystem in India

Event Readback

30th September 2024



## Table of Content

Introduction .....	3
Keynote by Dr. Sundar Narayanan .....	4
Panel discussion: Future of India's Responsible AI Ecosystem .....	5
Industry Showcase .....	7
Rocket Learning .....	7
OpenNyai - Jugalbandi .....	9
Wadhvani AI .....	11
Roundtable: All about the bias .....	13
Way Forward .....	14
Resources .....	14





## Introduction

At Aapti, we have been deeply immersed in research on artificial intelligence (AI) over the past two years, exploring the critical questions that drive debates around safety, trust, and innovation in the development of emerging technologies. Throughout this journey, we've consistently sought insights from others within the AI ecosystem, using their perspectives to guide and enrich our research. As part of our broader AI research initiative, supported by Google.org, we organised "Shaping AI Futures: Blueprints for the Responsible AI Ecosystem in India," which was held at The Oberoi Hotel in New Delhi. This event brought together a diverse group of stakeholders to engage in meaningful dialogue on the future of Responsible AI, and AI governance in India, and to examine [Aapti's framework on bias in AI systems](#).

The timing of this research and event is significant, as discussions surrounding AI regulation in India are reaching a critical juncture. Convenings like this are crucial in fostering a shared understanding and building a common taxonomy among the diverse actors involved across the AI value chain. Our goal was to not only stimulate conversations around AI governance but also to establish foundational principles that can guide future developments in the field.

The event was structured in three parts. The first half featured a keynote address and a panel discussion, setting the stage for an insightful dialogue on the state of AI governance in India. The second half showcased case studies from Industry and an interactive workshop focused on bias in AI systems. This workshop aimed to achieve three core objectives:

1. Understanding the AI value chain: Identifying the various stages where biases can be introduced into AI systems.
2. Validating and refining the techno-societal bias framework: Gathering feedback from stakeholders to enhance our emerging framework that addresses sources of bias in AI.
3. Developing strategies for bias mitigation: Exploring the "AI black box" phenomenon within the value chain to create effective strategies for identifying and mitigating biases.

By bringing together key players from government, industry, civil society, and academia, we aimed to catalyse actionable insights and foster collaboration toward a more inclusive and equitable AI ecosystem in India.



## Keynote by Dr. Sundar Narayanan

Aapti's flagship R-AI event was kicked off by Sundar Narayanan. With over 18 years of experience in ethics development and dissemination, including four years dedicated to AI ethics research, Sundar brings a wealth of knowledge to the table. His work with the Ministry of Telecommunications' TEC Fairness working group and his efforts in embedding Responsible AI into products are complemented by his unique approach to engagement—using humour, creativity, gamification, memes, art, and short stories to communicate the complexities that ethical principles are grounded in.

Below are some key insights from Sundar's keynote address:

- **Ethics as a dynamic framework:** Ethics and governance are not fixed constructs; they have always evolved in response to societal norms, shaped by the philosophies and ideologies of different groups throughout history. This same dynamic evolution is now unfolding with AI, where ethical considerations are continuously being redefined.
- **Proactive harm prevention:** A coordinated approach is essential not only to mitigate harms associated with AI but to prevent them entirely. At the core of this strategy is the need for widespread awareness and education, ensuring that all stakeholders understand the potential risks and responsibilities of AI.
- **Transparency and stakeholder inclusion:** Incorporating transparency, early disclosure, and the active involvement of diverse stakeholders throughout the AI development process adds tremendous value. It ensures that the Responsible AI journey remains accountable and inclusive at every stage of the value chain.
- **Environmental considerations around AI:** It's crucial to remain mindful of AI's environmental impact. Evaluating AI use cases through the lens of their environmental footprint should be a key consideration in deciding whether they move forward into production.
- **Skill development for Responsible AI:** Equipping AI developers, business leaders, and decision-makers with the necessary skills to build AI responsibly is vital. This empowerment can amplify efforts, accelerating both ethical AI adoption and its positive societal impact.



## Panel discussion: Future of India's Responsible AI Ecosystem

The panel discussion at Shaping AI Futures aimed to build a common goal for India's Responsible AI (R-AI) future. It brought together multiple perspectives to chart a path that addresses how India should approach the question of governance of AI, given its unique cultural, social, and economic landscape. The panellists included

1. Dr Sivaramakrishnan, a research advisor at the Centre for Responsible AI (CeRAI) at IIT Madras and a member of the Global Partnership on AI (GPAI)
2. Mihir Kulkarni, an ML scientist at Wadhvani AI, was also part of the panel. He contributed to the committee that framed India's first AI governance guidelines and recommendations.
3. Sowmya Karun, who leads Public Affairs for Google in India, added her experience in building data solutions for social good while collaborating with cross-functional teams to deploy them on the ground.

With Soujanya Sridharan from Aapti Institute moderating the discussion, the conversation explored how AI can address public problems but emphasised the need for a cautious approach, grounded in the principles of Responsible AI.

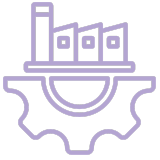
We believe the panel discussion was a pivotal conversation that highlighted the future possibilities of Responsible AI (R-AI) in India. It offered valuable insights into how we can shape this emerging dialogue around R-AI through multi-stakeholder collaboration and adaptive policy frameworks. The panel also explored the roles that various stakeholders—such as industries, CSOs, academia, and communities—can play in building AI systems that are not only beneficial but also incorporate mechanisms for responsibility and accountability to mitigate potential harms.

Below are some key points from the panellists, which give an important perspective of R-AI imagination for India,

- **Dr. Sivaramakrishnan** believed that a consultative and collaborative approach involving the right stakeholders is key. Looking at best practices from around the world, but also **building localised frameworks for regulating AI** which are outcome-oriented could be key pathways India takes in building an R-AI framework. Talking about what can be the regulatory focus, he opined that looking at sectoral use cases can be a better approach rather than focusing on the technology itself that evolves in dynamic, unpredictable ways.
- **Sowmya Karun** reiterated the need for a multistakeholder approach to framing responsible AI, **focusing on how industry players such as Google need to rely on other actors like CSOs and academia**. Further, a key challenge she points out is how to downstream AI principles. In this regard, she mentions how Google's approach to shared learning is by making its research, best practices, and toolkits openly

available, allowing actors access and opening up collaborative opportunities. Looking deeper into the problem of bias, she suggests a cautionary approach to what problems we aim to solve with AI. Given the complicated Indian diaspora, a measured approach, by building inclusive teams, many toolkits and attempts to understand the Indian landscape better (like Project Bindi) can help harness opportunities AI offers.

- **Mihir Kulkarni** shared his perspective and experience in building AI solutions for public health problems. **Prioritising agendas which are impact-focused and heuristics to identify where AI can help are the first steps to building value-driven AI systems.** Looking at fairness, he mentions how caste, religion and race in India work differently and thus there is a need to reorient such R-AI principles to specific contexts, articulating better problem statements. Connecting this thought to the development of AI, he mentions how there is a need to contextualise such problems to product development and machine learning, focusing on getting accurate data and using fair techniques. Lastly, looking at the roles various stakeholders play, he mentions three key agendas, protecting the rights of communities on the ground, having accountability watchdogs and realigning wrapped ideas of AI for the public.



## Industry Showcase

### Rocket Learning

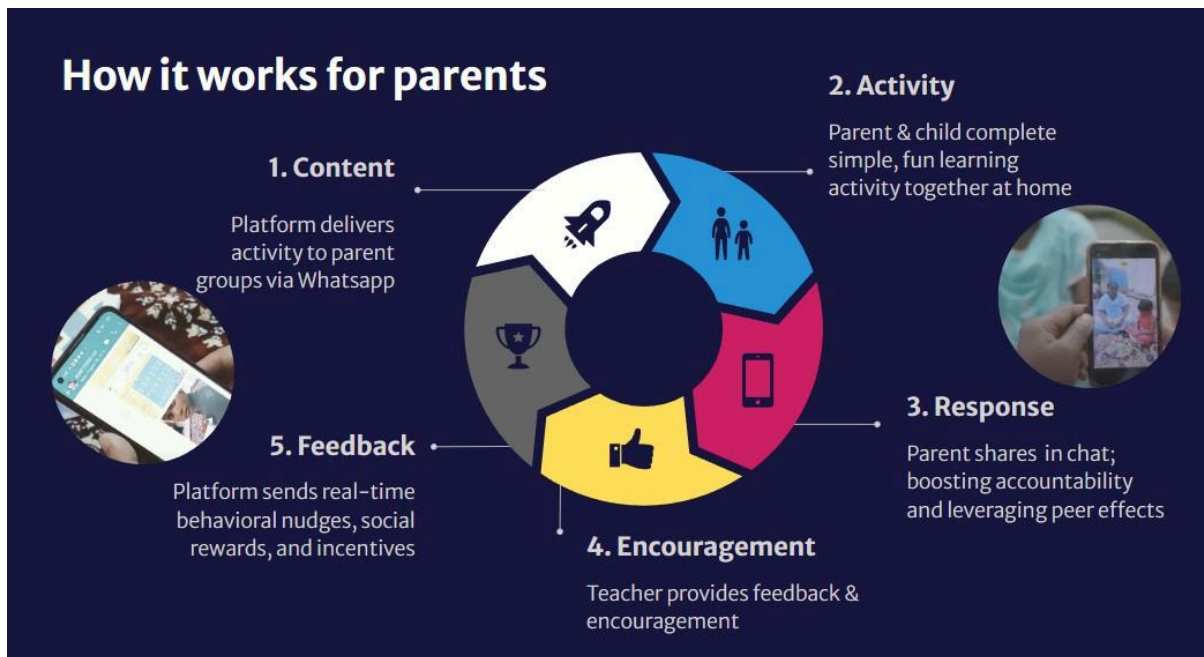
Akshat Goel is a specialist in machine learning and data science at RocketLearning, where he and his team are on a mission to bridge the learning gap for 75 million underserved preschool and primary school students across India.

Rocket Learning focuses on providing easily accessible, personalised homework content to young learners, particularly those from underserved communities. The platform empowers both parents and teachers—the key nurturing adults in a child’s early life—by giving them the tools and agency to actively engage in the child’s cognitive development. RocketLearning aims to prepare every child for both school and life, currently working with 2.2 million children in collaboration with nine state governments across India.

- The RocketLearning platform leverages WhatsApp to deliver engaging, age-appropriate learning activities directly to parents.
- Through these simple and fun activities, parents and children can learn together at home, promoting shared educational experiences.
- Parents are encouraged to share their participation in group chats, fostering accountability and utilising peer effects to motivate continued engagement.
- Teachers also play an active role, providing feedback and encouragement to parents, ensuring that the learning process is interactive and supportive.
- The platform enhances this engagement by delivering real-time behavioural nudges, offering social rewards, and using incentives to further drive participation and reinforce positive learning habits.



A few slides from RocketLearning's showcase:



## Goal 1: Automatic Homework Sheet Grading Deployed Successfully and Being Scaled

- **Developed internal data creation feature which allows us to label up to 4000 images per week, and used it to create a unique dataset of 60000 labeled examples covering 60 distinct worksheet types**
  - For training and fine-tuning purposes, we need a dataset of labelled worksheets. To create these labelled examples, we built a platform feature to synthesize labelled examples; our labelling team creates fresh examples on a rolling basis with this feature
- **Trained and fine-tuned a suite of openly available deep learning models for computer vision to greater than 90% accuracy on this dataset**
  - We treat this as an object detection problem rather than a classification problem; this is because we want to give children a detailed understanding of where they have been right and wrong, and not just a final grade; we have leveraged a variety of deep learning architectures, including pix2seq, a foundation model open sourced by Google Research
- **Built scalable infrastructure capable of processing > 50000 images per day in near real-time, including additional tools to filter out spam and identify which particular worksheet type is being processed**
  - We receive anywhere between 25,000 to 50,000 image interactions on our platform on a daily basis; not all of these images are worksheets; we deployed pre-trained OCR models to filter out non-worksheet interactions and also tag which particular worksheet types are being sent to us; this is also relevant for creating learning profiles
- **Conducted successful production deployment in two states and are now scaling up to different worksheet types and also to different geographies**
  - We have done a successful deployment in Haryana and Rajasthan, these geographies receive up to 8 different worksheet types a week; we have processed nearly 1.2 lakh images since this deployment began, for 18 different worksheet types on a rolling basis; by the end of March, we aim to scale up to all worksheets in our content bank and all geographies

## OpenNyai – Jugalbandi

Rashika Narain is actively involved in discovering entrepreneurs who are innovating at the intersection of law and technology. In addition to her work with OpenNyai, she closely collaborates with the sub-committee of the e-Committee of the Supreme Court of India, which aims to digitise and modernise India's legal infrastructure.

OpenNyai's Jugalbandi is an advanced AI-powered, multilingual, and multimodal conversational platform designed to provide information and enable large-scale citizen engagement. By harnessing cutting-edge AI, Jugalbandi empowers users with conversational AI solutions that function across various sectors and domains.

As a free and open platform, Jugalbandi integrates the capabilities of Generative AI with Indian language translation models developed under India's Bhashini Mission. This makes it possible to create powerful conversational AI tools tailored to diverse linguistic needs, ensuring accessibility across India's vast multilingual population.

The potential applications of Jugalbandi are vast and diverse, including:

- **Government Schemes:** Through collaboration with MyScheme.gov.in, Jugalbandi is already powering chatbots on WhatsApp and Telegram, allowing citizens to inquire about 121 government schemes in 10 different Indian languages. This democratises access to critical welfare information, making it easily available to those who need it most.
- **Law and Justice:** Applications built on the Jugalbandi stack, such as JIVA and JAGRIT, are designed to power chatbots and apps that can provide citizens with access to legal information. These tools can help break down barriers to legal literacy by delivering accessible, conversational AI-driven legal resources via WhatsApp and Telegram, ensuring more people can understand and engage with the justice system.
- **Healthcare:** Jugalbandi also has the potential to revolutionise healthcare access by powering WhatsApp and Telegram chatbots and apps that connect citizens with quality healthcare information and services. By bridging the gap between citizens and healthcare providers, these tools can help deliver timely, accurate medical information and facilitate healthcare access, particularly for underserved communities.

A few slides from OpenNyai's showcase:

# What is Jugalbandi

Jugalbandi is an advanced AI-powered multilingual/multi-modal conversational AI platform that can be used to provide information or enable action at scale for citizens.

jugalbandi

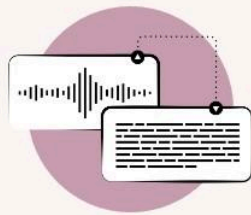
jugalbandi.ai / opennyai.org

## Salient Features



### Multi Lingual

Can be operated in multiple language, in particular 12 Indian languages supported by BHASHINI speech models.



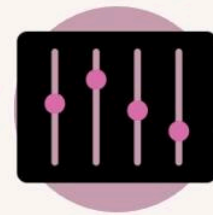
### Multi-Modal

Can both receive and send messages in both text and voice, so information still reaches people with low literacy.



### Scalable

Whether you have 100 users or 10,00,000 users, Jugalbandi is only limited by your infrastructure.



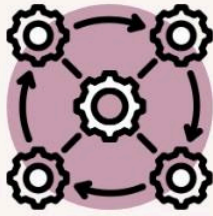
### Customisable

Customise every aspect of Jugalbandi, from which language models to use, to which LLM to use for reasoning

jugalbandi

jugalbandi.ai / opennyai.org

# Salient Features



## Interoperable

Works with common protocols and standards such as Beckn Protocol (PULSE, ONDC, OCEN, ONEST, etc.)



## Cloud Agnostic

Works on Azure, Google Cloud, and easy to customise for on-premise or other cloud platforms.



## Free/Open Source

Completely open-source code, built using Python and available on GitHub for anyone to deploy. Zero cost to test, using our hosted APIs.



## Easy to Deploy/Run

Get up and running within a day, and learn how to customise/manage the stack within 14 days. Get started on our free test API.

jugalbandi

jugalbandi.ai / opennyai.org

## Wadhvani AI

Mihir Kulkarni is a machine learning scientist at Wadhvani AI, where he focuses on developing AI and data science-driven interventions aimed at addressing pressing social issues. His work is part of Wadhvani AI's broader mission to leverage artificial intelligence for social good, particularly in underserved populations within developing countries.

The Wadhvani Institute for Artificial Intelligence is an independent, not-for-profit organisation committed to deploying AI solutions that benefit vulnerable and underserved communities. Their goal is to build scalable and sustainable systems by partnering with government bodies, non-profits, and other stakeholders to ensure their AI innovations reach those who need them most. Wadhvani AI's expertise spans critical domains, including healthcare, agriculture, education, and urban development.

Through a memorandum of understanding (MoU) with the government, Wadhvani AI collaborates to identify key use cases and problem statements aligned with national development priorities. This partnership grants access to restricted data sets and digital systems, enabling the team to build, deploy, and optimise AI models for public-sector platforms. The process involves various stages of the AI value chain, including compression, integration, evaluation, and iterative improvements. Ultimately, Wadhvani AI ensures knowledge transfer by providing training, tools, documentation, and support to ensure sustainable implementation.

Their diverse use cases demonstrate the real-world impact of AI in various sectors:

- Healthcare: Cough Against TB: An AI-powered application that analyses cough sounds and other symptoms to identify individuals with presumptive pulmonary tuberculosis (TB), facilitating early diagnosis.
- Agriculture: National Pest Surveillance System: A computer vision-based early warning system that helps farmers and agricultural experts identify and manage pests and diseases in key crops, providing timely advisories to protect national agricultural interests.
- Education: An AI-powered tool that offers automated insights from voice recordings of students reading passages. The system assesses reading comprehension and tracks learning and development indicators, helping educators monitor and enhance student outcomes.



## Roundtable: All about the bias

The roundtable focused on examining bias in AI systems and exploring ways to better understand and mitigate harmful biases. Aapti's research on AI bias served as the foundation for the discussion, using the value chain ontology to identify sources of bias and develop strategies for mitigation. The conversation centred around two key themes: first, defining what bias is and how we conceptualize it, and second, the operationalization of fairness in AI systems.

**Harmful vs. Acceptable Bias:** The debate centres around whether certain types of bias can actually be beneficial, especially in correcting historical inequities. This raises critical questions, how do we define harmful bias? And where do we draw the line between correcting for bias and introducing new biases? These considerations require nuanced thinking to avoid oversimplifying the conversation around bias in AI.

- *Bias as an amalgamation of the statistical concept in AI and ethical/ legal principles:* Biases within AI systems can be looked at as [statistical phenomena](#), which skew the results due to [external factors](#) (incorrect/unrepresentative training data, imperfect proxies, etc) making the values not representative of population realities. This understanding of bias itself is value-neutral and free from ethical judgments. The key is to identify arbitrary biases, which surmount harm and [discrimination](#).
- *Case studies as rubrics for understanding harmful biases:* To effectively address bias, we must look at real-world case studies. These provide insights into the degree of harm different biases can cause and reveal the complexities of mitigating them.
- *Working backwards from harm:* One effective approach to mitigating bias is to work backwards from harm. By identifying the potential negative impacts AI decisions can have on marginalised or vulnerable groups, we can trace those harms back to the specific biases embedded in the system.
- *Building a shared taxonomy across stakeholder groups:* A shared understanding of key concepts like bias, fairness, and responsibility is essential for effective collaboration across stakeholder groups. Establishing a common language—a shared taxonomy—ensures that these diverse groups can engage meaningfully in discussions around AI governance and ethical AI development.

**Fairness and Model Optimization:** Fairness is not just an ethical consideration but a key factor in optimising AI models. A fair model can lead to more robust and generalizable outcomes, improving performance across diverse populations.

- While fairness is often discussed from a moral or philosophical perspective, it's equally important to understand the mathematical and logical foundations that guide fairness in AI.

- Fairness measures, algorithms, and logic-based frameworks provide concrete tools for embedding fairness into AI systems. These structures allow us to systematically address bias and ensure that fairness is not just an abstract goal but a measurable outcome within the AI model itself.



## Way forward

- Based on the valuable feedback from the roundtable and the event we plan to revisit and refine our framework, which will be integrated into our larger report on bias in AI systems.
- Our report will also delve into the intersection of AI with digital integrity and cybersecurity. To further explore this, we hope to conduct another community of practice convening, soon.
- We greatly appreciate your insights and would love to continue the conversation. Please feel free to reach out to us at [contact@aapti.in](mailto:contact@aapti.in) or visit our website at <https://aapti.in/>.



## Resources

- [Rewiring Responsible AI: From Principles to Practice](#)
- [Decoding AI Development: The Efficacy of the Value Chain Ontology](#)
- [All About the Bias: Mapping and Mitigating Bias Across the AI Value Chain](#)