MATTER: Metrics and Assessment Tools for Trustworthy, Transparent, Explainable, and Reliable AI

Dr. Priyanka Narad*

Division of Development Research, Indian Council of Medical Research, New Delhi
Email: pnarad.hq@icmr.gov.in

Artificial Iintelligence's rapid deployment in domains such as healthcare, finance, and administration necessitates rigorous frameworks for trust, transparency, and explainability. **MATTER**—*Metrics and Assessment Tools for Trustworthy, Transparent, Explainable, and Reliable AI*—brings together cutting-edge benchmarks, interpretability methods, and governance practices for systematic evaluation. Benchmarks like **TrustLLM** and **AIR-Bench 2024** utilize 30+ datasets across 18 evaluation dimensions, including robustness to adversarial inputs, hallucination detection, privacy leakage analysis, fairness across subpopulations, and ethical reasoning. Technical advances in explainability are equally vital. Tools such as **LIME**, **SHAP**, and saliency mapping, embedded in **IBM's AI Explainability 360**, provide model-agnostic, human-understandable insights into decision pipelines. Governance metrics now integrate data lineage tracking, bias monitoring, and *adaptive trust calibration*, bridging technical evaluation with societal accountability. Despite progress, challenges remain in chain-of-thought reliability, dynamic bias detection, and ensuring fairness under domain shift—underscoring the need for holistic frameworks like MATTER.

From an Indian perspective, the **Indian Council of Medical Research (ICMR)** has pioneered AI ethics in healthcare through its *2023 Guidelines on AI in Biomedical Research and Healthcare*. These emphasize algorithmic transparency, human-in-the-loop validation, India-specific datasets, and task-based risk evaluation. ICMR is also investing in AI-driven clinical trial management, digital health registries, and context-sensitive bias assessments, aligning India's healthcare AI landscape with global best practices. In this talk, I will present the MATTER framework as a unifying architecture that bridges international benchmarks with ICMR's initiatives, enabling the development of AI systems that are technically robust, ethically sound, and societally aligned.

Keywords: MATTER framework, Trustworthy AI, Explainable AI (XAI), AI governance, Transparency, Reliability