

Annexure 3

NAME OF THE THEME	SELECTED APPLICANT	TITLE OF THE PROJECT	ABOUT THE PROJECT
Machine Unlearning	IIT Jodhpur	Machine Unlearning in Generative Foundation Models	To development novel method for targeted unlearning in open-source generative foundation models while minimising negative impact on overall model performance.
Synthetic Data Generation	IIT Roorkee	Design and Development of Method for Generating Synthetic Data for Mitigating Bias in Datasets; and Framework for Mitigating Bias in Machine Learning Pipeline for Responsible AI	To develop the algorithm and method for handling the bias at the model training and in-processing stage of ML model development
AI Bias Mitigation Strategy	National Institute of Technology Raipur	Development of Responsible Artificial Intelligence for Bias Mitigation in Health Care Systems	To develop responsible AI algorithms that reduce biases in medical system applications, image analysis, and diagnostic decisions
Explainable AI Framework	DIAT Pune and Mindgraph Technology Pvt. Ltd.	Enabling Explainable and Privacy Preserving AI for Security	Create AI models that provide accurate and interpretable results for human activity analysis for effective security in crowded environment.

Privacy Enhancing Strategy	IIT Delhi, IIIT Delhi, IIT Dharwad and Telecommunication Engineering Center (TEC)	Robust Privacy-Preserving Machine Learning Models	To develop robust distributed/federated learning algorithms which perform well in an adversarial environment susceptible to attacks
AI Ethical Certification Framework	IIIT Delhi and Telecommunication Engineering Center (TEC)	Tools for assessing fairness of AI model	To develop a three step certification process involving bias risk assessment, processing for metrics, and bias testing to ensure fairness for AI Systems in the Indian context
AI Algorithm Auditing Tool	Civic Data Labs	ParakhAI - An open-source framework and toolkit for Participatory Algorithmic Auditing	The proposed framework and toolkit will enable involving citizens in the responsible design, development, and deployment of algorithmic decision making systems.
AI Governance Testing Framework	Amrita Vishwa Vidyapeetham and Telecommunication Engineering Center (TEC)	Track-LLM, Transparency, Risk Assessment, Context & Knowledge for Large Language Models	To identify and address the specific gaps in the existing governance testing frameworks related to LLM;s downstream use-case and deployment
