Dr. Sandeep Kumar Satapathy

Associate Professor & Deputy Director,
Centre for Advanced Data Science,
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology, Chennai
Tamil Nadu, India

Email: sandeepkumar.s@vit.ac.in

Dr. Shruti Mishra
Associate Professor,
Centre for Advanced Data Science,
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology, Chennai,
Tamil Nadu, India

Email: shruti.mishra@vit.ac.in

Special Session on Data-Centric Artificial Intelligence: Foundations, Methods, and Real-World Applications

Artificial Intelligence (AI) has traditionally focused on developing increasingly sophisticated models. However, the paradigm is shifting toward Data-Centric AI, where the quality, diversity, and management of data are recognized as the primary drivers of AI performance. By emphasizing data collection, annotation, curation, augmentation, governance, and continuous improvement, Data-Centric AI aims to build robust, reliable, fair, and trustworthy AI systems that generalize effectively across diverse real-world scenarios.
This special issue invites original research, review articles, and application papers that advance the theory, methodologies, and practical deployment of Data-Centric AI across healthcare, finance, manufacturing, cybersecurity, transportation, smart cities, agriculture, education, and other domains.

Topics of Interest (including but not limited to):

- Data quality assessment, cleaning, and preprocessing for AI
- Data annotation, labeling strategies, and human-in-the-loop learning
- Synthetic data generation and data augmentation using generative AI
- Self-supervised, semi-supervised, weakly supervised, and active learning
- Data-centric foundation models and multimodal datasets
- Dataset bias detection, fairness, and responsible AI
- Data governance, provenance, privacy, and security
- Federated learning and privacy-preserving data management
- Benchmark dataset development, evaluation, and standardization
- Data versioning, lineage, and lifecycle management (DataOps/MLOps)
- Robust AI through data debugging, uncertainty estimation, and continual learning
- Knowledge graphs, metadata management, and data integration
- Explainable and trustworthy Data-Centric AI
- Domain-specific applications of Data-Centric AI in healthcare, climate science, agriculture, cybersecurity, robotics, and smart manufacturing.

We encourage contributions that demonstrate how high-quality, well-managed, and intelligently curated data can unlock the next generation of reliable, transparent, and scalable AI systems. This special issue aims to foster collaboration among researchers, practitioners, and industry experts working at the intersection of data engineering, machine learning, and responsible artificial intelligence.