

Global Dimensions of Multidisciplinary Research and Innovation



Editors

Dr. Neeta Bagalari

Dr. Tania Sur Roy

Dr. S. Valli Devasena

Ms. G. Brishya

Dr. E. Joseph Rubert

GLOBAL DIMENSIONS OF MULTIDISCIPLINARY RESEARCH AND INNOVATION

Dr. Neeta Baglari

Assistant Professor
Department of Education
Cotton University, Guwahati

Dr. Tania Sur Roy

Assistant Professor
Department of Education
Don Bosco University, Assam

Dr. S. Valli Devasena

Assistant Professor
Department of Commerce
Mother Teresa Women's University
Madurai, Tamilnadu

Ms. G. Brishya

Assistant Professor
Department of English
Agurchand Manmull Jain College
Meenambakkam, Chennai

Dr. E. Joseph Rubert

Professor
Department of Management Studies
Arunachala College of Engineering for Women
Manavilai, Kanyakumari District, Tamilnadu



www.merimaxpub.org

Edition: First

Year: November, 2025

ISBN: 978-81-995443-9-0

All Rights Reserved: No part of this publication can be stored in any retrieval system or reproduced in any form or by any means without the prior written permission of the publisher.

© **Publisher**

Publisher



Australia, Chennai & Kanyakumari

Phone: +91 6384730258

E-Mail: editor@merimaxpub.org

www.merimaxpub.org

PERFACE

. The accelerating complexity of the twenty-first century has placed unprecedented demands on the global research community. Challenges such as climate change, public health crises, technological disruptions, socioeconomic inequalities, and the transformation of educational and industrial landscapes require solutions that no single discipline can adequately provide. *Global Dimensions of Multidisciplinary Research and Innovation* emerges from this recognition, offering a platform that unites diverse fields of knowledge to explore new ways of understanding and addressing contemporary issues. This volume is founded on the belief that innovation thrives at the intersections of disciplines. When ideas, methodologies, and perspectives converge, they create fertile ground for groundbreaking discoveries and transformative practices. The contributions gathered here reflect the collaborative spirit of scholars, scientists, educators, industry leaders, and policymakers who share a commitment to expanding the boundaries of research. Their work illustrates how multidisciplinary approaches enrich inquiry, deepen analysis, and produce solutions that are more resilient, inclusive, and adaptable to global realities. In assembling this collection, we sought to highlight the varied yet interdependent domains that define modern research. The chapters encompass a wide range of subjects—science and technology, humanities and social sciences, business and economics, health and environmental studies, and emerging interdisciplinary fields. Whether examining digital transformation, sustainable development, cultural dynamics, or innovation ecosystems, each contribution demonstrates how knowledge flows across disciplines and geographical borders. Together, they reflect the interconnectedness of the contemporary world and the need for collaborative frameworks to navigate it effectively. A distinctive feature of this work is its international orientation. Contributors represent diverse cultural, academic, and institutional backgrounds, bringing global insights to local challenges and local perspectives to global debates. This plurality enriches the dialogue and underscores the importance of cross-cultural research partnerships in shaping the future of innovation. By weaving global voices into a unified narrative, the volume affirms that progress emerges not from isolated efforts but from collective engagement. We hope that this book serves as a valuable resource for researchers, students, practitioners, and policymakers seeking to broaden their understanding of multidisciplinary research. It is intended not only as a compilation of findings but also as an invitation—to question assumptions, to integrate new perspectives, and to participate in the ongoing evolution of knowledge across borders and boundaries. As the global community confronts ever more complex challenges, the insights presented here remind us that transformative ideas often arise where disciplines meet

Editors

EDITORIAL BOARD MEMBERS	
Dr. Neeta Baglari Assistant Professor Department of Education Cotton University, Guwahati Assam	Dr.Tania Sur Roy Assistant Professor Department of Education Don Bosco University, Assam
Dr.S.Valli Devasena Assistant Professor Department of Commerce Mother Teresa Women's University Madurai, Tamilnadu	Ms.G.Brishya Assistant Professor Department of English Agurchand Manmull Jain College Meenambakkam, Chennai
Dr.E.Joseph Rubert Professor, Department of Management Studies Arunachala College of Engineering for Women, Manavilai Kanyakumari, Tamilnadu, India	
Reviewer Committee Members	
Dr. Samson Davidson Dean-School of Computer Science and Information Technology University of Edenberg, Lusaka, Zambia	Prof. Dr. Marirajan Thiruppathi Deputy Vice Chancellor – Academic University of Edenberg Lusaka, Zambia
Dr. Velmurugan Thangarasu Director, Quality Assurance University of Edenberg Lusaka, Zambia	Mr.Alinani Simukanga Coordinator School of Computer Science and Information Technology University of Edenberg, Lusaka, Zambia
Dr.James Mumba School of Social Work University of Edenberg Lusaka, Zambia	Dr.Soniya K Associate Professor Department of Management St.Francis College, Bangalore
Dr. Cleopas P.J. Moono, School of Management Studies University of Edenberg Lusaka, Zambia	Dr.M.Priya Assistant Professor Department of Business Administration Nehru Arts and Science College Coimbatore
Dr.Thahira KK Assistant Professor Department of Economics Arignar Anna Govt Arts & Science College, Pondicherry	Dr.Sanitha KK Assistant Professor Department of Business Management Sree Narayanaguru Open University Kollam, Kerala

33	Sustainable and Circular Manufacturing in Mechanical Engineering <i>C. Anand Chairman, M. Pradeep , M. Suresh Kumar & S.R Sathish kumar</i>	254-260
34	Applied Data Science with Machine Learning and IoT: Big Data-Driven Innovations for Cross-Domain Research <i>Dr. Shanthi S</i>	261-265
35	Data Science at Scale: Big Data Architectures and Machine Learning Strategies for Real-World Insights and Multidisciplinary Innovation <i>Dr.S.Sridevi</i>	266-269
36	Blockchain-Enabled Secure Cloud Frameworks: Decentralizing Identity and Data Access Management <i>Ms. Battula Bhavya</i>	270-274
37	Cloud-Based Medical Imaging Platforms: Deploying Scalable Deep Learning Models for Disease Diagnosis (CT, MRI, X-ray) <i>Dr. Amirtha Preeya</i>	275-279
38	Cloud–Edge Collaborative Framework for Smart Mobility: Managing Traffic Data Streams Using Cloud-Based Analytics <i>Ms. Chaithra VS</i>	280-285
39	Computational Intelligence for IoT: Data Science and Big Data Analytics in Multidisciplinary Systems <i>Dr. Vinutha C B</i>	286-292
40	Data science at scale: Big Architecture and Machine Learning Strategies for Multidisciplinary Innovation and Real-World Research <i>Ms. S Thabassum Khan</i>	293-298
41	Medical Image Processing: Multidisciplinary Approaches for Clinical and Biomedical Research <i>Dr. Renukadevi</i>	299-304
42	Next-Generation Data Science: Machine Learning and Big Data Architectures for IoT and Multidisciplinary Research <i>Dr P Sudha</i>	305-309

Data Science at Scale: Big Data Architectures and Machine Learning Strategies for Real-World Insights and Multidisciplinary Innovation

Dr.S.SRIDEVI

Assistant Professor,
Department of CS with AI & ML,
NGM College, Pollachi

ABSTRACT

Data science at scale has emerged as a foundational driver of innovation across scientific, industrial, and societal domains. With the exponential expansion of big data from sensors, enterprise systems, cloud platforms, mobile devices, and scientific instrumentation, organizations require high-performance computational architectures and advanced machine learning strategies to transform raw data into actionable insights. This paper explores scalable data engineering pipelines, distributed data architectures, cloud-edge ecosystems, and large-scale learning models that enable the analysis of massive and heterogeneous datasets. It highlights the role of deep learning, graph analytics, automated machine learning (AutoML), and multimodal modeling in generating real-world impact. The study further examines multidisciplinary applications including healthcare, smart mobility, environmental science, industrial systems, finance, and social computing. Challenges involving data governance, bias, reproducibility, energy efficiency, and model interpretability are discussed, alongside future research pathways. Overall, data science at scale provides the backbone for developing intelligent systems capable of addressing complex real-world problems in an era defined by overwhelming data complexity and cross-domain integration.

KEYWORDS

Data Science, Big Data, Scalable Architectures, Machine Learning, Deep Learning, Distributed Systems, Cloud Computing, Real-World Analytics, Multidisciplinary Innovation, AI Systems.

1. Introduction

The rapid digitalization of industries, pervasive adoption of IoT devices, cloud-native ecosystems, and the integration of cyber-physical systems have led to unprecedented data growth. Organizations now handle terabytes to petabytes of real-time and historical data, requiring scalable infrastructures and efficient analytics workflows to extract meaningful insights [1]. Data science at scale addresses the unique computational demands posed by massive datasets through distributed processing engines, cloud-native architectures, parallelized algorithms, and advanced storage frameworks [2]. Machine learning models must be adapted for high-dimensional, high-velocity, and multimodal data, necessitating GPU acceleration, cluster-based training, and streaming analytics [3]. By combining big data engineering with intelligent learning models, scalable data science enables critical innovations in fields such as healthcare diagnostics, smart transportation, climate analytics, manufacturing automation, and financial risk modeling [4]. This synergy positions scalable data science as a critical enabler of real-world impact and multidisciplinary research.

2. Big Data Architectures for Scalable Insights

Modern big data ecosystems rely on distributed and fault-tolerant architectures capable of ingesting, storing, and processing massive datasets while ensuring low latency and high throughput. Distributed file systems such as Hadoop HDFS and cloud-native object storage solutions provide scalable storage architectures designed for high availability and parallel data access [5]. Batch processing engines like Apache Spark enable in-memory analytics and iterative computations, supporting complex transformations and ML training at scale [6]. For real-time applications, streaming engines including Kafka, Flink, and Spark Streaming facilitate low-latency event processing essential for predictive systems, intrusion detection, and sensor-driven analytics [7]. NoSQL systems—such as Cassandra, MongoDB, and HBase—provide schema flexibility and horizontal scalability for handling unstructured and semi-structured data [8]. Additionally, serverless computing and container orchestration platforms such as Kubernetes support elastic scaling, automated deployment, and microservices-driven analytics pipelines. These architectures allow organizations to operationalize complex ML workflows while maintaining reliability and cost efficiency.

3. Machine Learning Strategies for Large-Scale Analytics

Scaling machine learning requires computational optimizations, distributed training techniques, and model adaptation strategies that align with big data characteristics. Traditional ML algorithms must be parallelized to handle high-dimensional datasets, often using frameworks like MLlib, TensorFlow Distributed, and PyTorch Lightning [9]. Deep learning architectures—including CNNs, transformers, recurrent models, and graph neural networks—leverage GPU and TPU acceleration to support large-scale training over millions of samples [10]. Reinforcement learning algorithms are increasingly used in complex decision environments such as logistics optimization, traffic control, and industrial automation [11]. AutoML platforms automate feature engineering, hyperparameter tuning, and model selection, enabling non-experts to deploy ML models efficiently at scale [12]. Moreover, multimodal learning, which integrates text, images, time-series, and sensor data, supports real-world systems requiring contextual decision-making such as medical diagnostics and autonomous vehicles. These model strategies form the backbone of scalable AI designed for multidisciplinary domains.

4. Scalable Data Engineering and Cloud–Edge Ecosystems

Cloud computing plays a central role in enabling large-scale data science through elastic infrastructure, managed ML services, and distributed compute clusters [13]. Platforms such as AWS, Azure, and Google Cloud support end-to-end ML lifecycle management—from data ingestion and transformation to model training, deployment, and monitoring. Cloud-native architectures allow scalable data pipelines that integrate storage layers, ETL workflows, model registries, and monitoring dashboards. At the same time, edge computing provides local intelligence for low-latency applications involving autonomous vehicles, smart manufacturing, and healthcare monitoring [14]. Hybrid cloud–edge architectures distribute computational workloads, allowing edge nodes to preprocess data, perform micro-inference, and reduce bandwidth demands, while cloud layers handle large-scale model training and long-term analytics [15]. This collaborative environment supports real-time responsiveness,

resource optimization, and secure data management required in multidisciplinary applications.

. Multidisciplinary Applications of Scalable Data Science

5.1 Healthcare and Biomedical Analytics

Large-scale electronic health records, genomic data, medical imaging, and wearable sensor streams require advanced ML models capable of understanding complex biomedical patterns. Scalable CNNs and transformers are used for diagnosing diseases from CT, MRI, and X-ray images, while predictive models analyze patient risk factors and clinical outcomes [16]. Data science at scale supports personalized medicine, outbreak prediction, and real-time remote patient monitoring.

5.2 Smart Cities and Transportation Systems

Massive mobility datasets from GPS traces, traffic cameras, IoT sensors, and public transit systems enable scalable analytics for congestion prediction, route optimization, and adaptive signal control. Graph neural networks and LSTM models support large-scale traffic forecasting while reinforcement learning optimizes mobility flows in real time [17].

5.3 Climate Science and Environmental Monitoring

Environmental research uses petabyte-scale datasets from satellites, air quality sensors, and weather models. Scalable ML supports climate trend analysis, extreme event prediction, and environmental risk assessment, helping governments and organizations make data-driven policy decisions [18].

5.4 Industrial IoT and Manufacturing Automation

Industry 4.0 systems integrate data from sensors, robots, and machinery to enhance predictive maintenance, fault detection, and supply chain optimization. Digital twins combine simulation and real-time IoT data to create scalable models of manufacturing systems capable of predicting failures and improving operational efficiency [19].

5.5 Finance, Retail, and Social Systems

In finance, scalable ML supports fraud detection, credit scoring, and market trend forecasting using billions of transactions. Retail analytics rely on clustering, demand forecasting, and customer behavior analysis to optimize pricing and inventory strategies [20]. Social systems utilize large-scale sentiment analysis, misinformation detection, and behavioral modeling.

6. Challenges and Opportunities in Scalable Data Science

Despite its potential, data science at scale presents multiple challenges. The heterogeneity of multimodal datasets complicates preprocessing, integration, and quality assurance, necessitating advanced data validation tools and metadata management techniques [21]. Ethical challenges arise concerning privacy, consent, and responsible AI deployment, especially when working with healthcare, mobility, social, and financial data [22]. Large ML models require significant energy resources, prompting research into green AI, model

compression, and efficient hardware acceleration [23]. Ensuring fairness, transparency, and interpretability remains critical as systems influence high-stakes decisions in healthcare, policing, hiring, and finance. Reproducibility challenges persist due to evolving datasets and versioning complexities, highlighting the need for standardized data governance frameworks [24]. These challenges also present opportunities for future research in efficient algorithms, trustworthy AI, federated analytics, and decentralized intelligence.

7. Conclusion

Data science at scale has become a transformative foundation for multidisciplinary innovation, enabling organizations to derive value from massive, heterogeneous, and real-time datasets. By integrating big data architectures, scalable ML models, cloud–edge computing, and domain-specific analytics, data science empowers intelligent systems across healthcare, transportation, industrial automation, environmental monitoring, finance, and social platforms. As real-world systems continue to generate data at unprecedented rates, scalable data engineering pipelines, GPU-accelerated model training, and distributed AI infrastructures will become increasingly vital. However, challenges related to data quality, fairness, governance, privacy, and energy consumption must be addressed to ensure responsible and sustainable use. The future of data science lies in developing explainable, energy-efficient, privacy-preserving, and highly scalable AI systems capable of supporting global-scale innovation and solving complex societal problems. Ultimately, data science at scale acts as a catalyst for real-world insights, scientific advancement, and cross-domain transformation.

REFERENCES

- [1] Manyika, J. et al., “Big Data: The Next Frontier for Innovation,” McKinsey Global Institute, 2011.
- [2] Hashem, I. et al., “The Rise of Big Data on Cloud Computing,” Information Systems, 2015.
- [3] Dean, J., “Large-Scale Deep Learning,” Google Research, 2016.
- [4] Aggarwal, C. “Machine Learning for Big Data,” Springer, 2018.
- [5] Shvachko, K. et al., “HDFS Architecture,” IEEE MSST, 2010.
- [6] Zaharia, M. et al., “Spark: Fast and General Cluster Computing,” USENIX NSDI, 2012.
- [7] Kreps, J., “Kafka: Distributed Streaming Platform,” LinkedIn Engineering, 2015.
- [8] Stonebraker, M., “SQL vs NoSQL,” Communications of the ACM, 2012.
- [9] Meng, X. et al., “MLlib: Scalable Machine Learning,” JMLR, 2016.
- [10] LeCun, Y., Hinton, G., Bengio, Y., “Deep Learning,” Nature, 2015.
- [11] Sutton, R., Barto, A., “Reinforcement Learning,” MIT Press, 2018.
- [12] Feurer, M. et al., “AutoML: Automated ML Frameworks,” JMLR, 2019.
- [13] Mell, P., Grance, T., “NIST Cloud Computing Definition,” 2011.
- [14] Shi, W. et al., “Edge Computing: Vision and Challenges,” IEEE IoT Journal, 2016.
- [15] Cisco, “Fog Computing and IoT Architecture,” White Paper, 2015.
- [16] Litjens, G. et al., “Deep Learning for Medical Image Analysis,” 2017.
- [17] Li, Y. et al., “Diffusion Convolutional Neural Networks for Traffic Forecasting,” ICLR, 2018.





MERIMAX
INTERNATIONAL PUBLICATIONS

MERIMAX INTERNATIONAL PUBLICATIONS

Registered under Ministry of MSME, Govt of India UDYAM –TN-09-0080630

Kanyakumari, Tamilnadu, India (Branch: Australia & Chennai)

editor@merimaxpub.org | www.merimaxpub.org

PUBLICATION CERTIFICATE

This Publication Certificate has been issued to

Dr.S.SRIDEVI

Assistant Professor,
Department of CS with AI & ML,
NGM College, Pollachi

Publication of Book Chapter Entitled

**Data Science at Scale: Big Data Architectures and Machine Learning Strategies for
Real-World Insights and Multidisciplinary Innovation**

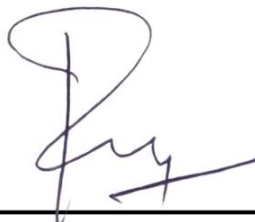
Published in ISBN Book Chapter

GLOBAL DIMENSIONS OF MULTIDISCIPLINARY RESEARCH AND INNOVATION

ISBN: 978-81-995443-9-0

Publication Date: November, 2025





DIRECTOR