Namit Shrivastava

College Park, MD | (240)-476-8513 | namit507@gmail.com
LinkedIn | GitHub

RESEARCH SUMMARY

Graduate researcher with proven expertise in survey methodology, causal inference, and advanced statistical modeling. Currently conducting geospatial epidemiological research at the University of Michigan's Institute for Social Research, developing stratified analytical frameworks for large-scale census tract data with complex missing data patterns. Published survey research in peer-reviewed venues (Springer) and presented at AAPOR on transformer-based sentiment analysis methodologies. Academic background combines rigorous statistical training with practical survey research experience, achieving Dean's Fellowship recognition. Research agenda focuses on developing trustworthy data integration frameworks that fuse survey methodology principles with generative AI, deep learning architectures, and privacy-preserving computational methods to advance automated data collection, quality assurance, and responsible measurement at scale.

EDUCATION

University of Maryland, College Park — Aug 2024 - May 2026
Master of Science, Survey & Data Science (Data Science Track) — College Park, MD

- **GPA: 3.792/4.0** | Dean's Fellowship Award Winner AY 2025-26
- Coursework: Statistical Modeling and Machine Learning I II, Fundamentals of Data Collection I II, Experimental Design and Causal Inference, Fundamentals of Computational Data Display, Long-Context Language Models, Machine Learning for Social Science, Applied Sampling

Birla Institute of Technology and Science (BITS) Pilani — Nov 2020 - Jul 2024
Bachelor of Engineering (Honours), Civil Engineering (Minor: Data Science) — Pilani, RJ, India

- **GPA: 7.93/10.0**
- Coursework: Foundations Of Data Science, Object Oriented Programming, Applied Statistical Methods, Fluid Mechanics, Construction Plan Tech, Analysis of Structures, Artificial Intelligence, Human Resource Development, Technical Report Writing, Computer Programming, Principles of Management, Electrical Sciences, Machine Learning, Mathematics (I, II and III), Operating Systems, Probability Statistics, Business Communication, Engineering Graphics, Data Mining, Data Visualization

PROFESSIONAL EXPERIENCE

Institute for Social Research, University of Michigan — May 2025 - Present
Research Assistant — Ann Arbor, MI, USA

- Architected and deployed a production-scale geospatial data integration pipeline processing **129,572 U.S. census tracts** across **3 RUCA** (Rural-Urban Commuting Area) strata, achieving **100% broadband data completeness** through multi-source fusion (FCC, ACS, CDC) and reducing baseline missingness by **28.6%** via composite measure imputation.
- Developed and validated a stratified epidemiological modeling framework analyzing COVID-19 inci-

dence patterns across **2,788 census tracts** with **100% social determinant coverage**, uncovering statistically significant non-linear rural inflections ($\mathbf{p < 0.01}$) and heterogeneous broadband effects by urbanicity.
- Designed and implemented a multi-stage data quality assurance protocol integrating Moran's I spatial autocorrelation analysis (identifying **15% spatial clustering violations**), temporal coverage validation (achieving **85% of tracts with 95% completeness**), and RUCA-stratified missingness diagnostics.

University of Maryland Feb 2025 - Jul 2025
Teaching and Graduate Assistant College Park, MD, USA

- Assisted Dr. Jörg Drechsler from Institute for Employment Research in teaching SURV735, effectively guiding **23 students** in understanding data privacy and confidentiality principles.
- Redesigned and standardized Canvas LMS infrastructure for **10+ JPSM instructors** supporting **125+ graduate students** by implementing modular course templates, automated grading workflows, and accessibility-compliant materials. This increased course satisfaction scores by **30%** and reduced instructor setup time by **40%**.

Legistify Services Private Limited Jan 2024 - Jun 2024
Machine Learning Engineer Gurugram, HR, India

- Engineered a scalable logo similarity detection system processing **2.4 million images** for intellectual property infringement analysis. Implemented perceptual hashing (imagehash), feature extraction (SIFT), and approximate nearest neighbor search (Faiss) via FastAPI microservice. Achieved **92% precision** on trademark conflict identification and reduced manual review time by **70%**.
- Deployed production Azure Cognitive Services API leveraging Vision Studio 4.0 for multi-modal document analysis (OCR, object detection, image captioning). Processed **50,000+ legal filings** with **95% extraction accuracy** and enabled bilingual text recognition (English/Hindi) using EasyOCR. This system directly supports **150+ client cases monthly**.
- Developed phonetic trademark similarity algorithm using double metaphone and Jellyfish string matching to compare **50,000 new trademarks** against **300,000 existing database entries**. Achieved **92% accuracy** in conflict detection while reducing computational overhead by **60%** through intelligent prefiltering and optimized SQL queries.
- Integrated and stress-tested Legistrak API ecosystem by executing **150+ REST API calls** via Postman with **95% success rate**. Identified **12 critical bottlenecks** and implemented caching strategies that increased data retrieval efficiency by **30%**. This improved platform responsiveness for **500+ daily active users**.

Accenture Jun 2023 - Aug 2023
Advanced Application Engineering Analyst Bangalore, KA, India

- Monitored enterprise security infrastructure by analyzing threat intelligence feeds from MITRE ATT&CK, AlienVault OTX, and proprietary sources. Achieved **89% accuracy** in threat classification and enabled proactive defense against **25+ emerging vulnerabilities**. This contributed to a **20% reduction** in incident detection latency.
- Conducted penetration testing and vulnerability assessments on **50+ web applications**, networks, and cloud infrastructure using Burp Suite, Nmap, and Metasploit. Identified and documented **120+ security vulnerabilities** (**15 critical, 45 high-severity**) that directly informed remediation roadmaps for Fortune 500 clients.
- Supported incident response operations during **3 security breaches** by performing log analysis with

Kusto Query Language (KQL) and Azure Sentinel. Assisted with containment and forensics while contributing to post-incident reports. This reduced mean time to recovery (MTTR) by **80%** through automated runbook implementation.
- Developed Python automation scripts for log parsing, threat indicator extraction, and security metrics dashboarding. Processed **10M+ daily log events** and generated executive-level threat summary reports that enabled data-driven security decision-making for C-suite stakeholders.

Indian Red Cross Society May 2022 - Jul 2022
Web Developer Bangalore, KA, India

- Designed and deployed a Drupal-based content management system for volunteer registry and donor database management. Streamlined volunteer onboarding workflows and reduced manual data entry time by **50%**.
- Localized IRCS main website by implementing multilingual support for Hindi and Kannada regional languages. Conducted comprehensive testing for functionality, performance, and security. Achieved **95% bug-free** user experience across **10,000+ monthly visitors**.

PUBLICATIONS AND PROJECTS

Achieving Sustainability in Supply Chain during Disruption Times: Role of Industry 4.0
Advances in Data-Driven Computing and Intelligent Systems (Springer) Aug 2023 - Feb 2024
BITS Goa, India | [Publication Link]

- Designed and executed a qualitative survey study with **200+ automotive supply chain experts** across India, employing fuzzy-set Analytical Hierarchy Process (FAHP) to rank **5 Industry 4.0 technologies** and **6 green supply chain practices**. Achieved consistency ratios below **0.10 threshold**, validating expert judgment reliability across all matrices.
- Developed a hierarchical Interpretive Structural Modeling (ISM) framework integrating MICMAC analysis to map causal relationships between digital technologies (IoT, Big Data Analytics, Cloud Computing) and sustainable practices. Identified government regulation and top management commitment as key drivers with highest driving power (score of **9 out of 9**).
- Established quantitative linkage between Industry 4.0 adoption and green supply chain performance through three-level hierarchical structure. Demonstrated that IoT, BDA, and cloud computing function as critical linkage factors (**driving power 7, dependence power 7**), enabling effective supplier-customer collaboration and circular economy practices.

Analyzing Public Sentiments of EVs in the Era of Sustainable Transformation
80th AAPOR Annual Conference Nov 2024 - May 2025
St. Louis, MO | [Conference Paper]

- Engineered a multi-source sentiment analysis pipeline processing **1.1M+ social media posts** from **5 Reddit communities** (**550K comments**) and **40 New York Times articles**. Implemented DistilBERT transformer model achieving **91.6% classification accuracy** on electric vehicle discourse spanning 2020-2024.
- Discovered systematic bias in large language model sentiment prediction through comparative analysis of **10 Groq LLM variants** (Llama 3.1/3.2 series, Mixtral). Found LLMs exhibited **+0.57 positive sentiment bias** compared to actual Reddit data (**M=-0.18, SD=0.52**), with statistical significance (**F(2,549)=28.43**).
- Identified temporal trends showing **35% increase in negative EV sentiment** on Reddit post-2022 despite rising adoption rates, while NYT coverage remained consistently negative (**M=-0.23**) for core EV terms but positive for Tesla and autonomous keywords. Demonstrated that Llama-3.2-90b-vision-

preview best approximated human sentiment patterns.

Transformer-Based Innovations in Internet Traffic Analysis and Scalable Energy Forecasting
Undergraduate Research Project BITS Pilani, Pilani, RJ, India

- Deployed and optimized advanced transformer-based models, achieving **97.8% classification accuracy** in internet traffic analysis and **98% reliability** in solar energy prediction.
- Demonstrated scalable solutions with **97% model size reduction** for efficient traffic classification and improved solar energy forecasting accuracy by **18%** for **500K+ data points**, while maintaining **99.9% uptime** under real-world conditions and handling over **10K+ requests/hour**.

TECHNICAL SKILLS

Programming: Python, R, Java, C++, JavaScript, TypeScript, HTML/CSS, Bash/Shell
Data & Databases: SQL (MySQL, PostgreSQL), NoSQL (MongoDB, Cassandra), Snowflake, Neo4j, Pinecone, Apache Spark, Kafka
AI/ML: PyTorch, TensorFlow, Keras, Hugging Face, LangChain, LlamaIndex, LangGraph, PySpark, NLTK, Spacy, SkLearn
DevOps: Git, Docker, Kubernetes, Jenkins CI/CD, REST APIs, Django, Flask, Angular, React.js, Terraform, Ansible
Cloud: AWS (EKS, S3, EC2, Lambda, Sagemaker), Azure, GCP, Tableau, Power BI, MLFlow, Qualtrics, STATA
Core: LLMs, Generative AI, RLHF/DPO, Deep Learning, NLP, Computer Vision, Survey Methodology, Causal Inference, MLOps

LEADERSHIP & SERVICE

Terrapin Leadership Institute, University of Maryland Aug 2024 - May 2025
Member College Park, MD

- Achieved **100% participation** in workshops on leadership, ethics, inclusion, and resilience
- Enhanced group dialogue quality by contributing reflective insights in all sessions, fostering more inclusive and engaging discussions among peers.

National Service Scheme (NSS), BITS Pilani Mar 2022 - Dec 2023
Executive Committee & Blood Donation Camp Core Team Member Pilani, RJ, India

- Planned **10+ activities** to improve English skills and confidence in villages around campus
- Organized blood donation camp coordinating **60+ volunteers**, managing data for **1,000+ donors**, achieving **844 successful donations** within 2 days

Peer Mentorship Program (PMP), BITS Pilani Aug 2021 - Dec 2023
Mentor Pilani, RJ, India

- Assisted Juniors with advice and materials to ensure they have a smooth transition into college life and succeed academically.

CERTIFICATIONS & ACHIEVEMENTS

Certifications: Microsoft Certified: Azure AI Fundamentals | API-based Programming (Postman) | AI For Everyone (Coursera) | Deep Learning & Machine Learning (Smartknower, Internshala)

Achievements: **1st rank** in Human Resource Development (**180 students**, BITS Pilani) | **Top 10** in Water & Wastewater Treatment (**98 students**, BITS Pilani) | JPSM Dean's Fellowship Award (AY 2025-26, UMD)