Navya Martin Kollapally
Curriculum Vitae

Email: nmartink@kean.edu
GitHub: https://github.com/navya777
LinkedIn: https://www.linkedin.com/in/navyamartin/
ORCID: https://orcid.org/0000-0003-4004-6508
Google Scholar: https://scholar.google.com/citations?hl=en&user=fr79IYkAAAAJ

Overview

Researcher and academic specializing in ethical AI, clinical informatics, and ontology development. My work seeks to integrate large language models into health data systems to improve fairness, transparency, and decision-making in clinical environments. My research is cross-disciplinary in nature in the domains of artificial intelligence, Natural language Processing, and deep neural-network optimization, with academic, government, and industry partnerships.

Education

New Jersey Institute of Technology (NJIT), Ph.D., Computer Science **August 2024**

Thesis title: A Methodological Framework for Ontology Development, Enrichment, and Application in Natural Language Processing Tasks

New Jersey Institute of Technology, MS., Computer Science **May 2021**

Mahatma Gandhi University, B.Tech., Electrical& Electronics Engineering **August 2012**

Awards and Honors

- Gonfalon Carrier for NJIT Ph.D. Commencement Ceremony representing graduating doctoral class in official processional, **May 2025**
- Best Reviewer award for American Medical Informatics Association (AMIA) Informatics Summit, **April 2024**
- Nominated for Excellence in Teaching (EiT) Award at NJIT, **April 2023**
- Best Student Paper Award at Knowledge engineering and Ontology Management (KEOD) conference, **May 2021**.
- Star Employee of Quarter at Infosys, **June 2014.**
- Innovative Research Award at ADSET (Accelerated Dissemination of Solar energy) INDO-GERMAN Dialogue, **February 2008**.

Academic Appointment

Kean University
Assistant Professor **August 2024- Present**

- Teaching two undergraduate courses: Database Management Systems and Java Programming
- Mentoring six undergraduate students as part of the Improving Undergraduate STEM Education (IUSE) initiative
- Member of the CAHSI (Computing Alliance of Hispanic-Serving Institutions) Mentor Network, guiding undergraduate and graduate research in Computer Science
- Serving on the Curriculum Committee for Fall 2025 academic planning

Doctoral Researcher **September 2021-May 2024**
Department of computer science, New Jersey Institute of Technology
Advisors: Dr James Geller and Dr Yehoshua Perl

- Devised an ontology tailored for practitioners to document injuries specifically pertinent to individuals of color.
- Formulated assessment methods to ensure the quality and coherence of ontological structures.
- Contributed to the development of a cardiology interface terminology, enabling fast skimming of unstructured Electronic Health Records (EHR) for physicians.
- Established a hyperparameter optimization framework to discern social context within EHR data.
- Utilized BIOE tagging to train ClinicalBioBERT sentence transformer to identify new concepts relevant to Cardiology Inference terminology from clinical notes.

New Jersey Institute of Technology **September 2021 – August 2024**
Teaching Assistant

- Served as a TA for Data mining, Introduction to Data Science, Advanced Database Management System, Database Management system and Foundations of Computer Science.
- Taught Introduction to Python for students with no coding background, motivating them to learn from errors and

assisting them to ace in programming.
- Tutored colleagues in data structure and algorithm for Ph.D. qualifying exam.
- Hosted office hours as additional support to clarify topics in "Introduction to Data Science class."

Blissful Coding Club **September 2021 – December 2022**

Advanced Python Instructor

- Created engaging lesson plans and hands-on coding projects to enhance students' understanding of advanced Python concepts.
- Conducted regular coding sessions, workshops, and code reviews, promoting a collaborative and inclusive atmosphere.
- Involve in maintenance of curriculum material and learning tool paddlet.

Industry Experience

Nokia Bell Lab, New Jersey **June 2024- August 2024**

Data Science AI/ML Intern

- Pending Patency application for the development of knowledge graph from technical specification using Llama3
- Developed an automatic pipeline for domain specific knowledge graph in telecommunication domain.

Department of Defense (DoD) **June 2022 – September 2022**

NLP researcher

- Conducted advanced data analytics such as delay estimation, delay prediction for ships under the regional maintenance center.
- Executed NLP analysis using Bert and LSTM on the contractor notes from maintenance dictionary.
- Studied delay real-time monitoring for naval ships, which belong to data management for AI problems.
- Developed effective strategies to deal with wide data problem in ML.

Infosys Ltd, Texas **October 2018 – December 2019**

QA Team Lead

- Lead and guide an offshore team of 7 members in projects involving compass application used by salespeople at Tiffany store.
- Closely interacted with Business team at Tiffany & Co. in leading compass projects and coordinating maintenance activities.
- Explained technical requirements to non-technical audience like third party app vendors.
- Pooled resources with a team of 6 members in offshore setting: delegated tasks, monitored progress, and maintained project deadlines.
- Identified the ETL jobs which was hindering transaction flow from POS terminal to lower QA environments, this issue was a bottle neck in many projects in pipeline accounting for a major loss in profit margin.
- Supervised all functions that reduce defect leakage in production.

Infosys Ltd Mysore, India **October 2012 – May 2014**

Test Engineer

- 2 years of experience in Software Quality Assurance in Banking and Financial sector
- Involved in testing on various levels for SunTrust Bank, Atlanta.
- Involved in maintaining the software quality of Teller Assist application and other financial software in SunTrust Bank.

Scholarship

Journal articles

1. N. M. Kollapally, J. Geller, V. K. Keloth, Z. He, and J. Xu, "Ontology enrichment using a large language model: Applying lexical, semantic, and knowledge network-based similarity for concept placement," Journal of Biomedical Informatics, p. 104865, 2025. Ranking- **Q1** (Scimago - Computer Scienec and Application) **H-index:137**
2. N. M. Kollapally, J. Geller, P. Morreale, and D. Kwak, "An Ontology for Social Determinants of Education (SDoEd) Based on Human-AI Collaborative Approach," Journal of Computing. Science in College., vol. 40, no. 3, pp. 191–203, 2024. Ranking: H-index:39

Conference proceedings

1. N. M. Kollapally and J. Geller, "Safeguarding Ethical AI: Detecting Potentially Sensitive Data Re-Identification and Generation of Misleading or Abusive Content from Quantized Large Language Models," in BIOSTEC (2), 2024, pp. 554-561. Ranking- H-index:12, acceptance ratio- 26%
2. N. M. Kollapally and J. Geller, "Hyperparameter Optimization Using Genetic Algorithm for Extracting Social Determinants of Health Text," in BIOSTEC (2), 2024, pp. 300-307. Ranking- H-index:12, acceptance ratio- 26%

3. N. M. Kollapally, M. K. H. Dehkordi, Y. Perl, J. Geller, F. P. Deek, H. Liu et al., "Using clinical entity recognition for curating an interface terminology to aid fast skimming of EHRs," in 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2024: IEEE, pp. 6427-6434. Ranking- H-index:25, acceptance ration-19.3%
4. N. M. Kollapally, V. K. Keloth, J. Xu, and J. Geller, "Integrating commercial and social determinants of health: A unified ontology for non-clinical determinants of health," in AMIA Annual Symposium Proceedings, 2023, vol. 2023, pp. 446.Ranking-Q1:Health Informatics. H-index:72, acceptance ratio-22%
5. N. M. Kollapally, Y. Chen, J. Xu, and J. Geller, "An ontology for the social determinants of health domain," in 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2022: IEEE, pp. 2403-2410. Ranking- H-index:25, acceptance ration-19.3%
6. M. K. H. Dehkordi, N. M. Kollapally, Y. Perl, J. Geller, F. P. Deek, H. Liu et al., "Skimming of Electronic Health Records Highlighted by an Interface Terminology Curated with Machine Learning Mining," in BIOSTEC (2), 2024, pp. 498-505 Ranking; H-index:12, acceptance ration-19.8%
7. N. M. Kollapally, Y. Chen, and J. Geller, "Health Ontology for Minority Equity (HOME)," in KEOD, 2021, pp. 17-27.
8. J. Geller and N. M. Kollapally, "Detecting, reporting and alleviating racial biases in standardized medical terminologies and ontologies," in 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2021: IEEE, pp. 1-5. Ranking- H-index:25, acceptance ration-19.6%

Book Chapters

1. Li, C., Kollapally, N.M., Chun, S A., Geller, J. (2022) Fake News Detection and Behavioral Analysis: Case of COVID-19, Book chapter in M. Last, M. Litvak and L. Miao (eds.) Detecting Online Propaganda and Misinformation. World Scientific, Singapore.

Panelist

1. National Center for Women and Information Technology (NCWIT) Aspirations in Computing Award- Kean University, *2025*

Invited speaker

1. New Jersey Alliance for Clinical and Translational Science (NJ ACTS)-St. Peter University, 2024
   Topic: Social Determinants of Health in Electronic Health Record
2. Data Science Symposium – *New Jersey Institute of Technology (NJIT), 2023*
   Topic: Advance of Large Language Models in Medical Informatics

Grants

External grants

1. **NSF 2541097**: CAREER: Automatic Ontology Enrichment: Enabling Multi-Domain Semantic Interoperability- Submitted July 2025, status- In review
2. **NSF 2523525:** Social Determinants of Education: Developing integrated framework for emotion detection- Submitted December 2024, status- rejected (reason change of NSF priority)- in process for resubmission
   Co-PI: Madjiguene Fall, Ph.D. (Department of Education, Kean University, USA)
3. **CAHSI- Google Institutional Research 2024-2025**: Cognitive LLMs for personalized phishing training Submitted February 2025, status rejected.
   PI: Kuldeep Singh, Department of Computer Science, University of Texas, El Paso

Internal grants

1. KEAN Summer SPF: Misinformation detection LLM March 2025, status rejected.

Teaching and Mentoring

Courses taught

1. **Computer Programming (CPS 2231)**
   - Delivered hands-on curriculum in Java programming, integrating AI-assisted coding tools to improve debugging and logic-building skills.
   - Implemented pair programming and group problem-solving labs to foster collaboration and peer learning.
   - Earned strong student feedback, with average evaluation ratings exceeding 4.5/5(Spring 2025) for clarity, approachability, and course relevance
2. **Database System Concepts and Applications (TECH 3740)**
   - Developed and delivered curriculum covering relational database design, SQL querying, normalization, and transaction management with emphasis on real-world applications.
   - Integrated AI-based database design assistants and visualization tools to help students conceptualize and optimize ER diagrams and schema structures

- Achieved average student evaluation rating of 4.55/5 (Fall 2024), with positive feedback on practical skill development and industry relevance.

Mentoring/Co-advising

- Mentoring group of nine undergraduate students on various Natural Language Processing tasks
- Thesis committee of graduate student Maliha Haider as Co-advisor, thesis title: *A Sentiment-Aware Pipeline for Analyzing and Summarizing Customer Reviews,* May 2025
- Mentoring five undergraduate students as part of Improving Undergraduate STEM Education (IUE)
- Serving as a Computing Alliance of Hispanic-Serving Institutions (CAHSI) Affinity research group (ARG) Certified Mentor, guiding undergraduate research projects in computing disciplines

Service

Professional Service

1. **Journal Peer Reviewer**
   - Journal of the American Medical Informatics Association (JAMIA) — 2024–2025
   - PLOS ONE — 2025
2. **Conference Reviewer**
   - American Medical Informatics Association (AMIA) Annual Symposium — 2022–2025
   - International Conference on Knowledge Engineering and Ontology Development (KEOD) — 2022–2023
   - IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM) — 2024
   - IEEE International Conference on Healthcare Informatics (ICHI) — 2024
3. **NSF Panel Reviewer**
   - OMNI Panel, National Science Foundation — March 2025
4. **Working Group Leadership & Involvement**
   - Women in AMIA — Leadership and mentoring activities
   - Knowledge and Data Mining Working Group, AMIA — Leadership member
   - OHDSI NLP Working Group — Active contributor and collaborator in natural language processing initiatives within the OHDSI research community.
5. **Workshops and Professional development**
   - Ethics in AI-Northeaster university, August 2025
   - NSF CISE proposal development workshop by American Society for Engineering Education, October-December 2024
   - NSF CAREER workshop at Kean-McAllister and Quinn, June 2025

Department

- Represented Department of Computer Science at the Spring and Fall 2025 Graduate Open House, engaging with prospective students to promote graduate programs, highlight curriculum strengths, and answer program-specific inquiries.

Certifications

- Federal Aviation Academy-Part 107: Small Unmanned Aircraft Systems (March 2025-March 2027)
- Coursera- Deep Learning with PyTorch: Siamese Network
- Coursera- Fine Tune BERT for Text Classification with TensorFlow

Technical Skills

- **Programming Languages:** Python, Java, SQL, HTML/CSS, JavaScript
- **Data Science & Analysis:** Pandas, NumPy, SciPy, Scikit-learn, TensorFlow, Keras, Matplotlib, Seaborn
- **Machine Learning & AI:** Supervised/Unsupervised Learning, Deep Learning, NLP, Model Evaluation, Data Preprocessing
- **Database & Tools:** MySQL, MongoDB, Oracle DB
- **Cloud & DevOps:** AWS (EC2, S3), Google Cloud Platform, Git, GitHub
- **Data Visualization:** Matplotlib, Plotly
- **Web & App Development:** Node.js, Xcode
- **Tools & Platforms:** Jupyter Notebook, VS Code, Anaconda, MATLAB, PyCharm, Postman
- **Statistical Analysis:** Hypothesis Testing, Regression, Time Series Analysis