Terre Haute, IN
(202) 213-7301
kazinazmulhasan@gmail.com
scholar.google.com/citations?user=QPwHFmEAAAAJ
github.com/nazmulkazi
huggingface.co/nkazi
linkedin.com/in/nazmulkazi

Nazmul Kazi

Curriculum Vitae

A universe that is unknowable is no fit place for a thinking being. — Carl Sagan

Education

2022–2023 **Master of Science in Computer and Information Sciences**, GPA: 3.83
University of North Florida, Jacksonville, FL

Thesis: *Automated short-answer grading and misconception detection using large language models.*

2019–2021 **Master of Science in Computer Science**
Montana State University, Bozeman, MT

Thesis: *Automated clinical transcription for behavioral health clinicians.*

2013–2018 **Bachelor of Science in Computer Science**
Montana State University, Bozeman, MT

Certifications

2023–2026 **Group 2 Social Behavioral Research Investigators and Key Personnel**
CITI Program (required for IRB application/approval)

Research Interests

Deep Learning, Natural Language Processing (NLP), Large Language Models (LLMs), Parallel Computing, Quantum Computing, Networking, Distributed Systems

Skills

Programming: Python, Rust, C/C++, R, Java, LaTeX, BST, PHP, JavaScript, jQuery, HTML, CSS, SQL, TypeScript, Perl, Bash, Documentation

Web Frameworks: Laravel, Django, Vue.js, Bootstrap, Tailwind CSS

Miscellaneous: Git, Windows, Linux-based OS, Android, Raspberry Pi, JupyterLab, VS Code, TensorFlow, PyTorch, Docker, LXD

Soft Skills: Swift learner, Active listener, Critical thinker, Problem-solver, Researcher, Engaging presenter, Effective communicator, Team player, Leadership, Versatile, Collaboration, Emotional intelligence, Decision maker

Experiences

2024–Present **Data Scientist**, FruitVaccine Inc., Champaign, IL (Remote)

2022–2023 **Graduate Research Assistant**, School of Computing, University of North Florida, Jacksonville, FL

2021–2021 **Internship**, FruitVaccine Inc., Champaign, IL

2019–2021 **Graduate Research Assistant**, School of Computing, Montana State University, Bozeman, MT

2015–2018 **Automated Web-System Developer**, Mechanical Engineering Department, Montana State University, Bozeman, MT

2015–2018 **Computer Programmer**, Architecture Department, Montana State University, Bozeman, MT

2015–2015 **Web Application Developer**, Office of International Programs, Montana State University, Bozeman, MT

2015–2015 **Web Developer**, Political Science Department, Montana State University, Bozeman, MT

Services

Program committees: ACL 2024-Present; ICMLA 2023–Present; BioNLP 2022–Present

Honors & Awards

2021 Student Scholarship Award, Association for Computational Linguistics and International Joint Conference on Natural Language Processing (ACL-IJCNLP)

2021 Student Scholarship Award, North American Chapter of the Association for Computational Linguistics (NAACL).

2021 Professional Advancement Grant, Spring, Graduate School, Montana State University.

2020 **Graduate School Dean's Excellence Scholarship**, Fall, Montana State University

2020 Professional Advancement Grant, Summer, Graduate School, Montana State University.

2020 Intelligent Systems for Molecular Biology (ISMB) Fellowship Award, International Society for Computational Biology.

2019 West Coast NLP (WeCNLP) Summit Travel Grant, Facebook.

2018 **Presidential Emerging Scholar Award**, Montana State University

2018 Travel Grant to attend 11th International Conference on e-Learning and Innovative Pedagogies, Undergraduate Scholars Program, Montana State University.

2018 Travel Grant to attend National Conference on Undergraduate Research (NCUR), Undergraduate Scholars Program, Montana State University.

2018 Undergraduate Research Grant, Spring, Undergraduate Scholars Program, Montana State University.

Publications

2024 Guna Sekaran Jaganathan, **Nazmul Kazi**, Indika Kahanda, and Upulee Kanewala. Towards understanding root causes of real failures in healthcare machine learning applications. In: *2024 IEEE Conference on Software Testing, Verification and Validation (ICST)*, pages 430–433. IEEE, 2024.

Guna Sekaran Jaganathan, **Nazmul Kazi**, Indika Kahanda, and Upulee Kanewala. MLHCBugs: A framework to reproduce real faults in healthcare machine learning applications. In: *2024 IEEE Conference on Software Testing, Verification and Validation (ICST)*, pages 445–447. IEEE, 2024.

2023 **Nazmul Kazi**, Indika Kahanda, S. Indu Rupassara, and John W. Kindt Jr. Zero-shot information extraction with community-fine-tuned large language models from open-ended interview transcripts. In: *2023 International Conference on Machine Learning and Applications (ICMLA)*, pages 932–937. IEEE, 2023. DOI: 10.1109/ICMLA58977.2023.00138.

Nazmul Kazi and Indika Kahanda. Enhancing transfer learning of LLMs through fine-tuning on task-related corpora for automated short-answer grading. In: *2023 International Conference on Machine Learning and Applications (ICMLA)*, pages 1687–1691. IEEE, 2023. DOI: 10.1109/ICMLA58977.2023.00255.

Nazmul Kazi, John W. Kindt Jr, Indika Kahanda, Christopher da Costa, Robert Carnahan, Hugh Mason, Brenda A. Wilson, and S. Indu Rupassara. Perspective chapter: Natural adjuvants for mucosal vaccinesthe promise of tomatine as an inherent adjuvant in tomatoes. In: Prof. Mourad Aribi, editor. *New Topics in Vaccine Development*, chapter 2. IntechOpen, Rijeka, 2023. DOI: 10.5772/intechopen.112932.

Nazmul Kazi. Automated short-answer grading and misconception detection using large language models. Master's thesis, University of North Florida, Jacksonville, FL, December 2023.

John W. Kindt Jr, **Nazmul Kazi**, Indika Kahanda, Christopher da Costa, Robert Carnahan, Brenda A. Wilson, Hugh Mason, and S. Indu Rupassara. Perspective chapter: The most natural possible vaccine administered in the most natural possible way - noninvasive over injectable vaccine delivery routes. In: Prof. Mourad Aribi, editor. *New Topics in Vaccine Development*, chapter 19. IntechOpen, Rijeka, 2023. DOI: 10.5772/intechopen.112698.

Dawson Dodds, John W. Kindt Jr, Christopher da Costa, **Nazmul Kazi**, Joseph T. Mahoney, and S. Indu Rupassara. Supply chain logistics and business ecosystems needed for the development of natural vaccines with novel, safer, and noninvasive delivery mechanisms. In: Prof. Mourad Aribi, editor. *New Topics in Vaccine Development*, chapter 8. IntechOpen, Rijeka, 2023. DOI: 10.5772/intechopen.113953.

2022 S. Indu Rupassara, John W. Kindt Jr, **Nazmul Kazi**, and Indika Kahanda. Challenges and opportunities in current vaccine technology and administration: A comprehensive survey examining oral vaccine potential in the united states. *Human Vaccines & Immunotherapeutics*, 18(6):2114422, 2022. DOI: 10.1080/21645515.2022.2114422.

Nazmul Kazi. Automated clinical transcription for behavioral health clinicians. Master's thesis, Montana State University, Bozeman, MT, December 2022. `https://scholarworks.montana.edu/xmlui/handle/1/16632`.

2021 Srinivasan Sridhar, **Nazmul Kazi**, Indika Kahanda, and Bernadette McCrory. Psychiatry transcript annotation: Process study and improvements. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 10(1):71–75, 2021. DOI: 10.1177/2327857921101030.

Nazmul Kazi, Nathaniel Lane, and Indika Kahanda. Automatically cataloging scholarly articles using library of congress subject headings. In: Ionut-Teodor Sorodoc, Madhumita Sushil, Ece Takmaz, and Eneko Agirre, editors. *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop*, pages 43–49, Online, April 2021. Association for Computational Linguistics. DOI: 10.18653/v1/2021.eacl-srw.7.

James P Becker, Indika Kahanda, and **Nazmul Kazi**. WIP: Detection of student misconceptions of electrical circuit concepts in a short answer question using NLP. In: *2021 ASEE Virtual Annual Conference Content Access*, 2021. DOI: 10.18260/1-2–38076.

2019 **Nazmul Kazi** and Indika Kahanda. Automatically generating psychiatric case notes from digital transcripts of doctor-patient conversations. In: Anna Rumshisky, Kirk Roberts, Steven Bethard, and Tristan Naumann, editors. *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, pages 140–148, Minneapolis, Minnesota, USA, June 2019. Association for Computational Linguistics. DOI: 10.18653/v1/W19-1918.

Mohammad Anani, **Nazmul Kazi**, Matthew Kuntz, and Indika Kahanda. RDoC task at BioNLP-OST 2019: A mental health informatics task with research domain criteria. In: Kim Jin-Dong, Nédellec Claire, Bossy Robert, and Deléger Louise, editors. *Proceedings of the 5th Workshop on BioNLP Open Shared Tasks*, pages 216–226, Hong Kong, China, November 2019. Association for Computational Linguistics. DOI: 10.18653/v1/D19-5729.

Notable Projects

2022–2024 **Predicting circuit noise in near-term quantum devices using Deep Learning. Supervisor: Dr. Zornitza Prodanoff.**

Generated a comprehensive dataset comprising 240,000 pairs of 5-qubit quantum circuits with randomized gate placements. Executed each circuit 5,000 times on the IBM Burlington to capture mean output noise. Leveraging neural network to predict noise disparities between circuit pairs.

2022–2023 **Improving LLM performance on automated short-answer grading through transfer learning. Collaborators: Dr. Indika Kahanda.**

We explore and leverage the effectiveness of transfer learning on automated short-answer grading (ASAG) by fine-tuning the RoBERTa Large model on natural language inference (NLI) corpora for semantic inference prior to fine-tuning it on the task/target dataset, SciEntsBank. The model significantly benefits from the NLI corpora, particularly in enhancing its performance on the contradictory class which constitutes only 10% of the dataset. Transfer learning also shortens the performance gap between unseen answers and unseen questions/domains.

2021–2023 **Detecting student misconceptions in short answers using large language models. Supervisor: Dr. Indika Kahanda. Collaborators: Dr. James Becker, Dr. Corey Pittman, and Larry Snedden.**

Misconceptions, whether arising from misunderstood concepts or incomplete information, can create cognitive roadblocks that hinder the acquisition of accurate knowledge. When left unaddressed, these misconceptions tend to perpetuate, leading to a cascade of aftereffects that extend beyond the classroom. We developed a framework that leverages LLMs to detect common misconceptions prevalent among students enrolled in introductory STEM courses.

2021–2022 **Zero-shot information extraction with community-fine-tuned large language models. Collaborators: Dr. S. Indu Rupassara, John W. Kindt Jr., and Dr. Indika Kahanda.**

Machine learning tasks like data annotation, model training, and parameter tuning are resource-intensive and often limit their practicality for one-time data extraction, medium-sized datasets, or short-term projects. There are many community-fine-tuned large language models (CLLMs) that are fine-tuned on task-specific datasets and can demonstrate impressive performance on unseen data without further fine-tuning. Adopting a hybrid approach of leveraging CLLMs for rapid text data extraction and subsequently hand-curating the inaccurate outputs yields high-quality results, workload balance, and improved efficiency.

2018–2021 **Generating medical case notes from patient-provider conversations. Supervisor: Dr. Indika Kahanda & Dr. Matthew Kuntz. Collaborators: Dr. Eric Arzubi & Dr. Wade Hill.**

Electronic health records (EHRs) are notorious for reducing face-to-face time with patients while increasing screen time for clinicians leading to burnout. Generating the case notes automatically will provide clinicians with a means to spend more time in patient care than in EHRs.

2017-2018 **Developing an interactive cloud-based integrated learning environment for project-based hands-on compiler development. Collaborators: Dr. Indika Kahanda.**

An online cloud-based interactive coding ground intended for teaching a project-based compiler course. The system comes pre-installed with the required tools and modules to provide a ready-to-use platform for the students. It also incorporates instant auto-grading of their submissions using pre-defined test cases.

References

Dr. S. Indu Rupassara **indurupassara@gmail.com**
Co-Founder/CEO, FruitVaccine Inc.

Dr. Indika Kahanda **indika.kahanda@unf.edu**
Assistant Professor, School of Computing, University of North Florida

Dr. Zornitza Prodanoff **zprodano@unf.edu**
Professor/Interim Director, School of Computing, University of North Florida

Dr. Sourav Sen Choudhury **schoudhury@psiquantum.com**
Senior R&D Engineer, PsiQuantum