LEVI KASTER

k.levi@wustl.edu | https://www.linkedin.com/in/levikaster/ | St. Louis, MO

EDUCATION

Washington University School of Medicine in St. Louis — **St. Louis, MO**
PhD in Biomedical Informatics and Data Science — *August 2024 - Present*

- **Cumulative GPA:** 4.00/4.00
- **Advisors:** Philip R. O. Payne and Aditi Gupta

Washington University in St. Louis - McKelvey School of Engineering — **St. Louis, MO**
Bachelor of Science in Data Science — *August 2020 - December 2023*

- **Cumulative GPA:** 4.00/4.00 **| Valedictorian | Engineering Class Rank:** 1/452

JOURNAL MANUSCRIPTS

Published/Pre-Prints

Kaster, L., Hillis, E., Oh, I.Y. et al. Comparison of Rule- and Large Language Model-based Phenotype Extraction from Clinical Notes for Neurofibromatosis type 1 (NF1). Journal of The American Medical Informatics Association (2025). https://doi.org/10.1093/jamia/ocaf155

Kaster, L., Hillis, E., Oh, I.Y. et al. Automated extraction of functional biomarkers of verbal and ambulatory ability from multi-institutional clinical notes using large language models. J Neurodevelop Disord 17, 24 (2025). https://doi.org/10.1186/s11689-025-09612-w

Zhang, H., Xu, T., Cao, D. et al including **Kaster, L**. OmniCellTOSG: The First Cell Text-Omic Signaling Graphs Dataset for Joint LLM and GNN Modeling. arXiv preprint arXiv:2504.02148 (2025). https://doi.org/10.48550/arXiv.2504.02148

Baldridge, D.*, **Kaster, L.***, Sancimino, C. et al. The Brain Gene Registry: a data snapshot. J Neurodevelop Disord 16, 17 (2024). https://doi.org/10.1186/s11689-024-09530-3

** Shared first-authorship*

Submitted and Under Review

Krussel, A., Gupta, A., **Kaster, L.**, et al. (2023). MedMap: A usability test of an AI enhanced ontology mapping tool in the health sciences classroom. [Submitted for publication to Medical Science Educator by Springer]

POSTERS/ORAL PRESENTATIONS

Oral Presentations

Levi Kaster, Inez Oh, Casey Vickstrom, Virginia Lanzotti, Philip Payne, Christina Gurnett, Aditi Gupta (2024). Utilizing Large-Language Models to Extract Patient Verbal and Ambulatory Status from Multi-Institutional and Multi-Specialty Clinical Notes. Platform Presentation at 2024 Child Neurology Society Conference; November 2024, San Diego, CA

- *Submitted as an abstract, 7% (32/466) of which were accepted as oral presentations*

Levi Kaster, Rui Mu, Ethan Hillis, Inez Oh, Stephanie M. Morris, David H. Gutmann, Randi E. Foraker, Philip R. O. Payne, Aditi Gupta. (2023). A Text-Mining Model for Extracting Phenotypes from NF1 Clinical Notes. Lightning talk at I2DB Symposium; April 2024, St. Louis, MO

Levi Kaster, Rui Mu, Ethan Hillis, Inez Oh, Stephanie M. Morris, David H. Gutmann, Randi E. Foraker, Philip R. O. Payne, Aditi Gupta. (2023). A Text-Mining Model for Extracting Phenotypes from NF1 Clinical Notes. Midstates Consortium for Math and Science Undergraduate Research Symposium; November 2023, St. Louis, MO

Posters

Levi Kaster, et al. (2025). Comparing Prompting Methods for Entity Extraction from Clinical Notes with RAG and Divide-and-Conquer Novel Methods Development. To be Presented at American Medical Informatics Association (AMIA) 2025 Annual Symposium; November 2025; Atlanta, GA

Levi Kaster, et al. (2024). Identifying NF1 Phenotypes from Unstructured Clinical Notes Using a Text-Mining Based Phenotype Extraction Model. Presented at American Medical Informatics Association (AMIA) 2024 Informatics Summit; March 2024; Boston, MA

Levi Kaster, Inez Oh, Zachary Abrams, Phillip Payne, et al. (2023). A De-Identification Algorithm to Automatically Identify and Remove PHI from Structured Multi-Scale Data from the Brain Gene Registry. Presented at American Medical Informatics Association (AMIA) 2023 Informatics Summit; March 2023; Seattle, WA

Levi Kaster, Rui Mu, Ethan Hillis, Inez Oh, Stephanie M. Morris, David H. Gutmann, Randi E. Foraker, Philip R. O. Payne, Aditi Gupta. (2023). A Text-Mining Model for Extracting Phenotypes from NF1 Clinical Notes. Washington University Fall 2023 Undergraduate Research Symposium; October 2023; St. Louis, MO

Virginia Lanzotti, et al including **Levi Kaster**. (2023) The Brain Gene Registry: An ongoing, collaborative resource of the Intellectual and Developmental Disabilities Research Centers (IDDRC). Presented by Virginia Lanzotti at 2023 American Society of Human Genetics Annual Meeting; November 2023; Washington D.C.

Maya Chopra, Virginia Lanzotti, Inez Oh, Aditi Gupta, **Levi Kaster**, et al. (2023) Clinical Variants for Brain Gene Curation: A Powerful and Under-Utilized Resource. Presented at 2023 Gatlinburg Conference; April 2023; Kansas City, MO

RESEARCH HONORS

Top Prize in Predictive Analytics at the (CAIDF) Hackathon *Chicago, IL; 2025*

- Awarded $5,000 Prize
- Developed multi-institutional 30-day fall readmission risk models that incorporated LLM-extracted phenotypes from clinical notes
- Displayed re-admission risk though an interactive dashboard that incorporated patient level LLM-generated summaries, model determined re-admission risk, and the top risk factors.

3rd Place Prize in Risk Prediction Datathon at I2DB Annual Symposium *St. Louis, MO; 2025*

- Awarded Small Monetary Prize ($200)
- Provide a greater description

Award for Best Computer Science Research Paper (Out of >500 students in Technical Writing) *2023*

- "Is your Health Information Safe? A Review of Current De-identification Protocols"

EXPERIENCE

Biomedical Informatics & Data Science PhD Program in Division of Biology and Biomedical Sciences
Graduate Student Researcher *August 2024 - Present*

Institute for Informatics, Data Science and Biostatistics at Washington University School of Medicine
Bioinformatics Research Assistant (Full Time) *January 2024 - August 2024*
Research Assistant (Part-Time) *August 2022 - December 2023*
BIDS@I2DB Summer Research Internship *May 2022 - August 2022*

Hengen Lab at Washington University in St. Louis **St. Louis, MO**
Research Assistant *Sept. 2021 - May. 2022*

Subaru of Indiana **Lafayette, IN**
Data Analyst Intern *June 2021 - Aug. 2021*

TEACHING EXPERIENCE

Teaching Assistantships
Washington University School of Medicine in St. Louis
BMI 5201 Biomedical Data Science I Fall 2025

Bridge2AI - Voice Summer School
Artificial Intelligence and Machine Learning from Voice Data Class Summer 2024

Washington University in St. Louis - McKelvey School of Engineering
CSE 314 Data Manipulation and Management Spring 2023, Fall 2023
CSE 240 Logic and Discrete Mathematics Spring 2022, Fall 2022

ADDITIONAL PROJECTS

CSE 543T Project: Using Machine Learning to Predict NBA Single Season Statistics
GitHub Link: https://github.com/kasterlevi/CSE543_NBA_Project

- Compared linear regression, neural networks, k-NN, and neural network models for predicting nba player statistics with a group of 4. Culminated in a Research Paper.
- Skills: Webscraping (BeautifulSoup), Machine Learning (SkLearn), Data Analysis, Delegation

CSE 314A Project: A Kaggle Competition to Predict the Value of 6th Men in the NBA
GitHub Link: https://github.com/kasterlevi/6thManDataProject/tree/main

- Created a mock Kaggle competition as a part of a group of 3. Using a Dataset we constructed, participants try to predict a team's success from the statistics of their 6th man.
- Skills: Database construction, API Requests, Querying

ACADEMIC HONORS & SCHOLARSHIPS

- McKelvey School of Engineering Valedictorian *2024*
- Dean's List Every Semester *2020 - 2023*
- McKelvey Engineering Professional Development Stipend Award ($750) *2023*

- Antoinette Frances Dames Engineering Award *2022*
- Steven R. and Susan A. Lowy Scholarship *2021*
- Forum Credit Union Scholarship *2020*
- Robert J. Denari Science Scholarship *2020*

TECHNICAL SKILLS

Technical Skills: Python, R, SQL

Softwares/Tools: PyTorch, Transformers, Langchain, Ollama, OpenAI API, LSF job scheduler, Azure/Databricks Environment, AWS, Github, R-Markdown

Activities: Organizer for Genetics Student Run Seminar, Community Coordinator for Ultimate Frisbee Club, Member of Powerlifting Club, Varsity College Football (1 year)

REFERENCES AVAILABLE UPON REQUEST