Zachary Levonian

Computer Science PhD
zacharylevonian@gmail.com
Tacoma, WA, USA
+1 (608) 780 9028
https://levon003.github.io

Expertise: human-computer interaction, data science, natural language processing

Experience

- **Digital Harbor Foundation** — Baltimore, MD (Remote)
  Senior Machine Learning Engineer — *May 2023–Present*
  - Designed and implemented chatbot interfaces using large language models.
  - Released open-source Python library for retrieval-augmented generation.
- **ConcertAI** — Cambridge, MA (Remote)
  Senior Machine Learning Engineer — *Nov. 2022–Apr. 2023*
  - Deployed production ML models for patient ranking and missing data imputation.
  - Led an observability overhaul: data validation, new metrics, and integration tests.
- **CaringBridge** — Eagan, MN (Remote)
  Data Scientist Intern — *Jun. 2021–Aug. 2021*
  - Designed, implemented, and evaluated a recommendation system for health blogs.
  - Implemented first-in-org data and analysis pipelines for user search and follow data.
- **Amazon** — Seattle, WA (Remote)
  Applied Scientist Intern — *Sep. 2020–Dec. 2020*
  - Designed an interactive interface for data annotation using React and Flask.
  - Implemented active learning and full-text search methods using PyTorch and Lucene.
  - Compared sampling methods for interactive machine learning using simulations.
- **The MITRE Corporation** — McLean, VA
  Computer Scientist — *Jan. 2015–Jul. 2017*
  - Designed and developed automatic speech recognition (ASR) safety interfaces as researcher in the Center for Advanced Aviation System Development (CAASD).
  - Improved ASR performance on air traffic controller and pilot radio transmissions through the application of cutting-edge techniques from academia.
  - Architected and implemented a Hadoop-based capability for large-scale processing of air traffic controller radio transmissions data.
- **General Dynamics Mission Systems** — Fairfax, VA
  Software Developer — *Oct. 2014–Jan. 2015*
  - Provided design, integration, and software development support for research & development team using geospatial and graph databases for multi-modal data.

Publications (Refereed Conference and Journal)

- **Z. Levonian**, M. Zent, N. Nguyen, M. McNamara, L. Terveen, S. Yarosh, "Some other poor soul's problems': a peer recommendation intervention for health-related social support," *Proc. ACM Hum.-Comput. Interact.* and the Conference on Computer-Supported Cooperative Work (CSCW), 2025. arxiv.org/abs/2209.04973
  - Designed, implemented, and evaluated a recommender system for health blogs.
- **Z. Levonian**, O. Henkel, C. Li, M.-E. Postle, "Designing Safe and Relevant Generative Chats for Math Learning in Intelligent Tutoring Systems," *Journal of Educational Data Mining (JEDM)*, 2025. doi:10.5281/zenodo.14751365
  - Implemented retrieval-augmented generation for an education question-answer task.
  - Designed and executed usability survey of large language model conversation.
  - Previously accepted and presented at the Educational Data Mining (EDM) 2024 conference and at a NeurIPS 2023 workshop. See: arxiv.org/abs/2310.03184
- C.E. Smith, H. Miller Hillberg, **Z. Levonian**, "'Thoughts & Prayers' or '❤️ & 🙏': How the Release of New Reactions on CaringBridge Reshapes Supportive Communication in Health Crises," *26th ACM Conference on Computer-Supported Cooperative Work and Social Computing* (CSCW), 2023. doi:10.1145/3610035
  - Designed a user survey & conducted semi-structured interviews.
- **Z. Levonian**, M. Harper, C.-J. Lee, V. Murdock, "Trade-offs in Sampling and Search for Early-stage Interactive Text Classification," *27th Annual Conference on Intelligent User Interfaces* (IUI), 2022. doi:10.1145/3490099.3511134
  - Quantitative simulations of small-sample interactive machine learning.
  - Methods: ML classification, active learning, full-text search, NLP transformer models
- **Z. Levonian**, M. Dow, D. Erikson, S. Ghosh, H. Miller Hillberg, S. Narayanan, L. Terveen, S. Yarosh, "Patterns of Patient and Caregiver Mutual Support Connections in an Online Health Community," *23rd ACM Conference on Computer-Supported Cooperative Work and Social Computing* (CSCW), 2021. doi:10.1145/3434184
  - Quantitative social network analysis of an online health community.
  - Methods: regression, ML classification, social network analysis, content analysis
- **Z. Levonian**, D.R. Erikson, W. Luo, S. Narayanan, S. Rubya, P. Vachher, L. Terveen, S. Yarosh, "Bridging Qualitative and Quantitative Methods for User Modeling: Tracing Cancer Patient Behavior in an Online Health Community," *Proceedings of the 14th International AAAI Conference on Web and Social Media* (ICWSM), 2020. 10.1609/icwsm.v14i1.7310
  - Developed method to incorporate qualitative themes into quantitative user modeling.
  - Conducted a comparison of ML and keyword-based classification approaches.
- C.E. Smith, **Z. Levonian**, H. Ma, R. Giaquinto, G. Lein-Mcdonough, Z. Li, S. O'Conner-Von, S. Yarosh, "'I Cannot Do All of This Alone': Exploring Instrumental and Prayer Support in Online Health Communities," *ACM Transactions on Computer-Human Interaction* (TOCHI), 2020. doi:10.1145/3402855
  - Conducted statistical analysis on survey of patients and their support networks.

  - Methods: frequentist statistics, survey design, visualization

- H. Miller Hillberg, **Z. Levonian**, D. Kluver, L. Terveen, and B. Hecht, "What I See is What You Don't Get: The Effects of (Not) Seeing Emoji Rendering Differences across Platforms," *Computer Supported Cooperative Work* (CSCW), 2018.
  - Conducted statistical analyses of survey response data.
- S. Chen, H. D. Kopald, R. S. Chong, Y.-J. Wei, and **Z. Levonian**, "Readback Error Detection using Automatic Speech Recognition," *Air Traffic Management Research and Development Seminar 2017* (ATM), 2017.
  - Conducted literature review and drafted background on acoustic modeling for ASR.
- S. Chen, H. D. Kopald, A. Elessawy, **Z. Levonian**, and R. M. Tarakan, "Speech Inputs to Surface Safety Logic Systems," *IEEE/AIAA 34th Digital Avionics Systems Conference* (DASC), 2015.
  - Computed quantitative automatic speech recognition (ASR) results.

Publications (Poster & Workshop)

- **Z. Levonian** and O. Henkel, "Safe Generative Chats in a WhatsApp Intelligent Tutoring System," *Educational Data Mining (EDM) Workshop: Leveraging Large Language Models for Next Generation Educational Technologies*, 2024. `arxiv.org/abs/2407.04915`
- O. Henkel, L. Hills, A. Boxer, B. Roberts, **Z. Levonian**, "Can Large Language Models Make the Grade? An Empirical Study Evaluating LLMs Ability To Mark Short Answer Questions in K-12 Education," short paper at *Learning @ Scale*, 2024. `doi:10.1145/3657604.3664693`
- **Z. Levonian**, L. Hagen, L. Li, J. Lilleboe, S. Wastvedt, A. Halfaker, L. Terveen, "ORES-Inspect: A technology probe for machine learning audits on enwiki," *Wiki Workshop*, 2024. `arxiv.org/abs/2406.08453`
  - Implemented a React-based web interface for visualizing machine learning predictions.
- S. Bhandari, **Z. Levonian**, R. Annand, J.-Z. Baxter, J. Cain, J. Flach, J. Joseph-Thomas, J. Mueller, "Digital Trial Solutions eScreening: a software solution that ranks patients by their predicted clinical trial eligbility using Real-World Data and Machine Learning," abstract in the *American Society of Clinical Oncology Annual Meeting* (ASCO), 2023.
  - Analyzed system usage data and conducted semi-structured interviews.
- R. Wan, **Z. Levonian**, S. Yarosh, "How much is a 'like' worth? Engagement and retention in an online health community," poster in the *23rd ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW), 2020.
  - **Outstanding Poster Recognition – CSCW 2020** (Top 5% of accepted posters)
  - Fit survival analysis models to predict user retention from engagement data.
- M. Butzer, **Z. Levonian**, Y. Luo, K. Watson, Y. Yuan, C.E. Smith, S. Yarosh, "Grandtotem: Supporting International and Intergenerational Relationships," poster in the *23rd ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW), 2020.

  - Designed and developed a prototype to facilitate intergenerational communication.

- P. Vachher, **Z. Levonian**, H.-F. Cheng, S. Yarosh, "Understanding Community-Level Conflicts Through Reddit r/place," poster in the *23rd ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW), 2020.
  - Quantified conflicts using social media log data analysis.

- C. Li, **Z. Levonian**, H. Ma, S. Yarosh, "Condition Unknown: Predicting Patients' Health Conditions in an Online Health Community," poster in the *21st ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW), 2018.
  - Trained and evaluated text classifiers for online post authors' health conditions.

Education

- **PhD in Computer Science** — Minneapolis, MN
  University of Minnesota — *Sep. 2017–Oct. 2022*
  - Researched human-computer interaction (HCI) and social computing for health
  - Led quantitative and qualitative research teams
  - Advisors: Loren Terveen and Svetlana Yarosh
  - Coursework: HCI & UI Technology, Social Computing, Machine Learning, Social Network Analysis, Embodied Computing, Database Systems, Statistics & Regression

- **BA in Computer Science (*magna cum laude*)** — Northfield, MN
  Carleton College — *Sep. 2010–Jun. 2014*

Mentorship & Teaching

- **Research mentor of undergraduates** — Minneapolis, MN
  University of Minnesota ProDUCT Lab — *Sep. 2017–Aug. 2022*
  - Mentored **21** undergraduate and Master's students, with student work presented in 3 full conference papers, 3 conference posters, 2 BA honors theses, and 1 MS thesis.
  - 2020 CRA Outstanding Undergraduate Researcher Awards: Honorable Mention for my student Drew Erikson. (Honors thesis: topic modeling for health communities)
  - Mentor for UMN Human-Centered Computing REU 2021 and Big Data REU 2019.
  - Mentored topics include: quantitative modeling, topic modeling, propensity score matching, embedding models, survival analysis, data science, visualization.

- **TA for *CSCI1001: Overview of Computer Science*** — Minneapolis, MN
  University of Minnesota CS Department — *Jan. 2018–May 2018*
  - Managed five undergraduate TAs along with grading and lab section logistics.
  - Designed new unit and problem set on relational databases & SQL.
  - Maintained weekly office hours period for myself and other TAs.

- **Prefect for *CS202: Mathematics of Computer Science*** — Northfield, MN
  Carleton College CS Department — *Jan. 2014–Mar. 2014*
  - Organized weekly prefect sessions of 5-15 students to review material.

Research Funding & Awards

- **Doctoral Dissertation Fellowship** Minneapolis, MN
  University of Minnesota *Sep. 2021–May 2022*
  - 1-year University-level dissertation fellowship to study peer recommendation systems.
- **Research Assistant** Minneapolis, MN
  University of Minnesota CS Department *Sep. 2017–May 2021*
  - Researching social support, online communities, classification methods, recommendation systems, and value sensitive design.
- **Early Career Research Program funding recipient** McLean, VA
  $83,000 of Funding *Oct. 2016–Sep. 2017*
  - Proposed research applying contemporary semantic parsing techniques to transcriptions of air traffic controller radio transmissions in order to extract meaning.
  - Implemented and compared NLP techniques to an existing corpus of air traffic controller communications.
- **Distinction in integrative exercise (B.A. capstone)** Northfield, MN
  Carleton College CS Department *Sep. 2013–Mar. 2014*

Service & Other Experience

- **Graduate Research and Discussion Seminar Co-facilitator** Minneapolis, MN
  University of Minnesota *Jan. 2019–Apr. 2020*
  - Organized biweekly seminar discussion for graduate Computer Science researchers.
- **Reviewer**
  SIGCHI Conferences *Jan. 2019–Present*
  - Reviewed for ICWSM {2019, 2020, 2021, 2022, 2023}, CSCW {2020, 2021, 2022, 2023, 2024}, WWW {2021, 2022}, CHI {2022, 2024}, JEDM 2025, AIED 2025.
- **GroupLens Seminar Co-facilitator**
  University of Minnesota CS Department *Jan. 2020–Apr. 2021*
  - Recruited external speakers and organized research discussions for GroupLens.
- Other: REU Researcher (Montana State University), Lead Writing Consultant (Carleton College), Tutor and Lab Assistant (Carleton College), Lead Facilitator for Sexuality and Gender Activism group (Carleton College).

Programming Skills

In academic settings and in industry. Italics indicate less than one month of experience.

- Languages: Python, Java, R, C, Objective C, Scheme, Bash, SQL, JavaScript, HTML+CSS, *Visual Basic, Perl, Awk, x86 Assembly, C++, Clojure, C#, Lua, Groovy*
- Frameworks: PyTorch, SciPy, NumPy, Matplotlib, Pandas, scikit-learn, Hadoop MapReduce, Spark, SpaCy, React, AWS (S3, EC2), HuggingFace Transformers, Dash, lifelines, *fast.ai, NLTK, gensim, node.js, Lucene/Solr*

- Tools: Maven, Git, Mercurial, Vim, various Java and Python IDEs, Jupyter, LaTeX, SQLite, MySQL/MariaDB, Postgres/Redshift, Hive, Vowpal Wabbit, Docker, Slurm, Apache Oozie, *Splunk, Jenkins, Kubernetes, Make, MongoDB, Redis, Ant, CVS, Valgrind, Xcode*
- Operating Systems: Ubuntu, Fedora, Red Hat Enterprise Linux, Windows, Mac OS X
- GitHub: `https://github.com/levon003/`

Other Skills

- Statistics: Frequentist hypothesis testing, regression (linear, logistic, Poisson, softmax), classification (SVM, GBM, MLP), NLP (BoW, TF-IDF, embeddings, sequence tagging), deep learning (optimization, regularization), evaluation, Bayesian modeling (basics)
- User Research: experiment design, A/B testing, contextual inquiry, interviews, surveys, usability studies, log/exploratory data analysis, content analysis, user stories.

Other Materials

- I've write blog posts: `https://levon003.github.io`
- I gave talks at conferences (CSCW, ICWSM, IUI, EDM), at companies (Amazon, CaringBridge, MITRE, LEVI), and at graduate seminars.