David Shumway

PhD Student

Email: davidshumway@gmail.com
Github: @davidshumway
LinkedIn: @david-shumway-0b4661159

Phone: (773) 759-5970
Address: 160 N. Elizabeth St., Unit 1909, Chicago, IL 60607

Professional Experience

Graduate Research & Teaching Assistant, University of Illinois at Chicago, Chicago, IL *2018-Present*

- Worked with state-of-the-art methods related to knowledge graphs, ontologies, AI, GIS, ETL, data science, visual analytics, SWfMSs, community web portals, machine learning, transfer learning, domain adaptation, LLMs, and spatiotemporal datasets towards solving real-world problems.
- Taught curriculum to 5-250 students (11 semesters) and led weekly discussions with 60-90 students (4 semesters) for various courses including Languages & Automata, Software Development for Mobile, Data & Web Semantics, Big Data Mining, Digital Literacy, Software Engineering I, and Software Design.
- Instructed 15 students in the graduate course Data & Web Semantics (Fall 2021).
- Mentored groups and individual undergraduate students in 4 semesters as part of NSF REU.

Co-Owner & Lead Developer, Finger Lakes Business Systems, Plattsburgh, NY *2010-2018*

- Collaborated with 750+ clients through Amazon MTurk, a competitive fast-paced on-demand virtual marketplace. Completed 1000+ unique jobs and 3+ million microtasks across various industries with a 99.9% task approval rating. Provided long-term solutions outside of the site with key clients.
- Designed solutions focused on web scraping, workflow optimization, data management, distributed systems, and API development, and built various methods to streamline the Amazon MTurk worker interface. Wrote 375K+ lines of in-house code (JS, AJAX, HTML/CSS, LAMP, Bash, Python, cURL).

Summer Intern, École Polytechnique Fédérale de Lausanne, DSIL, Lausanne, Switzerland *Summer 2017*

- Assisted in building out Neo4j language translation database for Kamusi by integrating new datasets.

Editor & Tech. Specialist, Paul Brunton Philosophic Foundation (Burdett, NY) *2010-2012*

- Curated 100K+ scanned documents and provided LAMP, JS/HTML, Bash, VBA, and VB.NET solutions.
- Automated scan processes, resulting in 4x speed gain and saving upwards of 2 years' worth of work.

Various roles, Tech. Freelancer, Delivery Driver, Retail Associate, Library Clerk *2000-2010*

Education

Ph.D. in Computer Science, University of Illinois at Chicago (expected May 2025) *2018-Present*

- Advisors: Isabel Cruz (in memoriam), Cornelia Caragea.
- Completed PhD coursework requirement (Spring 2021) and PhD qualifier exam (2020, *Storage & Querying Provenance in Scientific Workflow Management Systems*).
- Notebook Reviewer: 2nd and 3rd Annual EarthCube Call for Notebooks (2021, 2022).

B.S. in Computer Science, State University of New York at Plattsburgh *2015-2017*

- SUNY Plattsburgh Dean's List (Fall '15, Spring '16, Fall '16, Spring '17, Fall '17).
- SUNY Plattsburgh Annual CS Department Academic Excellence Award (2015-2016).

Publications: **Knowledge Graph for Water Quality and Health Risks**, Gautam et al. (2023, co-author).

Example Projects

Microbial Risks to Divers at Coastal Locations (2021-Present)

- Developed scuba diving microbial risk ontology, and aided in use of LLMs toward ontology learning.
- Used tools such as Python, Jupyter, Google Colab, Pandas, GeoPandas, Matplotlib, scikit-learn, PyTorch, TF, and ADAPT for spatiotemporal dataset ETL; to build and evaluate supervised baseline, semi-supervised, and unsupervised time series regression models to predict microbial concentrations (e.g. RF, XGBoost, MLP, CNN, LSTM, GNN, EasyAdapt, subspace alignment, CORAL, MixUp); to analyze data and results; and to experiment with state-of-the-art methods in this domain.

MALDI-DB (2020-2021 - Django, Nginx, PostgreSQL, D3.js, Docker, Docker compose, R, RPlumber)

- Built Django public web portal hosted on Mass Open Cloud for analysis of bacteria and small molecules MALDI-TOF spectra, and integrated an R library for spectra processing via RPlumber.

Zoomie: Round-robin breakout scheduling in Zoom (2020-2021 - Firefox, Chrome)

- Designed extension utilizing Zoom's web client for groups interested in round-robin breakout rooms.

Chicago Schools App (2018 - D3.js, Leaflet.js, HTML/CSS)

- Developed 3 out of 5 visualizations, search menu, and single page functionality with linked views.

Tools for Amazon's Mechanical Turk (2013-2018 - Firefox, Chrome)

- Designed extension to improve worker productivity on the Amazon MTurk website (1k-5k weekly users).

Tech. Skills: Python, Java, PHP, HTML/CSS, JavaScript, C++/C#, R, PostgreSQL/MySQL, Android, Bash, Linux, Git, Pandas, GeoPandas, Jupyter, Google Colab, Matplotlib, scikit-learn, PyTorch, TensorFlow, Django, Docker, Mass Open Cloud, LaTeX, OSS, Protégé, RDF/OWL, REST, AJAX, cURL, Browser extensions.