Jason Portenoy

jason.portenoy@gmail.com | jasport.org

Education

University of Washington, Information School   *Seattle, WA*

PhD, Information Science   2021

- [Dissertation](): "Harnessing Scholarly Literature as Data to Curate, Explore, and Evaluate Scientific Research"

MS, Information Science   2017

Brown University   *Providence, RI*

BS, Neuroscience   2008

Work Experience

University of Washington   *Seattle, WA*

Data Engineer — Center for an Informed Public (CIP)   March 2022 -- Present

- Manage and maintain the data infrastructure supporting the research and operations of the CIP, an interdisciplinary center for research and action around misinformation and disinformation.
- Responsibilities include oversight and management of a collection of physical and virtual servers, and pipelines and tools for data collection, storage, and analysis.
- Apply updates and security patches, build and maintain containers (Docker and Kubernetes) for various tasks, monitor and manage a Kubernetes cluster which runs most of the computing operations of the Center.
- Set up and manage multiple databases with billions of rows and dozens of terabytes of data.

Research Analyst   Sept 2021 -- March 2022

- With the Gordon and Betty Moore Foundation, developed grantexplorer.org — a tool to explore federal research grants over time — using React, D3, and FastAPI for interactive visualizations, and Elasticsearch and Gensim language models to intelligently assist with keyword queries.

Research Assistant / Teaching Assistant   Sept 2013 -- June 2021

- Research projects with Center for an Informed Public, Military Suicide Research Consortium, National Academy of Sciences, JSTOR, Science History Institute, Pew Charitable Trusts
- TA experience: Advanced Data Science Methods, Calling BS in the Age of Big Data, Design Thinking, Information Systems Analysis and Management, Client Side Web Development

Allen Institute for AI (AI2)   *Seattle, WA*

Intern / Visiting Researcher — Semantic Scholar Team   March 2020 -- Present

- Built AI-enabled scholarly recommendation systems and wrote research publications.
- Using Pytorch and additional Python libraries, I fine-tuned a pretrained BERT language model to encode document similarity, which I used to create embeddings for 3.5 million documents.
- I built a custom React and D3 based web app to provide recommendations for users and collect data for user studies.
- Currently, I am employed part-time working to deploy these methods in a production setting.

Chan Zuckerberg Initiative   *Redwood City, CA*

Visiting Researcher   May 2019 -- Sept 2019

- Analyzed networks of biomedical researchers to improve suggestions of new papers in feeds.

Microsoft Research and AI   *Redmond, WA*

Summer Research Intern (Mentors: Paul Bennett, Ryen White, Eric Horvitz)   June 2017 -- Sept 2017

- Developed methods to identify when people make commitments in their written communications, and understand what kind of commitments they make. Co-authored a patent based on this work.
- Used email data from Microsoft's internal terabyte-scale Cosmos data store to build ML classifiers

Skills

Python, Pandas, JavaScript, React, Typescript, R, SQL, D3, HTML, CSS, Linux, Git

Network Analysis, Visualization, High Performance Computing, Machine Learning and AI, Databases

For publications and additional info, see [full CV]()