

Alyssa Smith

Curriculum Vitae

May 2022

Address: 294 Windsor St Apt 1, Cambridge, MA 02139, USA
Phone: +1 401 688 2239
Email: asmith99@gmail.com
Website: lentil-soup.github.io
[LinkedIn](#) — [Twitter](#) — [Github](#)

Education and Qualifications

- 2013 - 2017 **BS** in Humanities and Engineering (21E) with Comparative Media Studies and Computer Science **Massachusetts Institute of Technology**
- 2021- present **PhD** in Network Science; **Advisors:** Prof. Brooke Foucault Welles and Prof. David Lazer **Northeastern University**

Professional Experience

- 2021-present **Research Assistant** I've been doing a mix of data science and data engineering in my research, as well as some critical work on social movements; right now I'm studying a few aspects of online amplification and movement-building. **Network Science Institute, Northeastern University**
- 2019-2021 **Labs Data Scientist** I worked creatively and iteratively to design new models that impact user experience on the KAYAK website, with a great deal of freedom in model choice, design process, and evaluation; I collaborated with product managers and software engineers regularly to ensure that the project will be useful to stakeholders. I also was a co-lead of the LGBTQ+ Employee Resource Group. **KAYAK**
- 2017-2019 **Graph Data Engineer** Built and maintained models; handled data ingestion, pipelining, and testing; and designed services. Worked with diverse and often unpredictable data sources, and therefore had to create extremely robust, adaptable pipelines. Also worked on improving diversity and inclusion in the workplace. **Kensho Technologies**
- 2016 **Product Analytics Intern** Produced actionable insights through gathering and analysis of user data. Created a neural network language model of user-created tags and clustered users based on site usage data. **Tumblr**
- 2015 **Intern** Assessed the efficacy of municipal broadband programs in Massachusetts; published a MediaCloud blog post analyzing linking behavior in media coverage of the GamerGate controversy. **Berkman-Klein Center for Internet and Society, Harvard University**

Skills

- Elasticsearch
- SQL-style databases
- Graph Databases (GraphQL)
- data ETL
- h5py, hdfs
- Git
- AWS CLI
- Unix Command Line
- Pytorch
- {sci,num}py/pandas
- sklearn/nltk
- NLP
- Bayesian Inference
- Simulated Annealing
- Ant Colony Optimization
- gephi/igraph/networkx
- Apache Airflow
- Redis
- Web Scraping
- Twitter API
- Python 2.x, 3.x
- Java
- R

Coursework

Undergraduate

- 6.046 Design and Analysis of Algorithms
- 6.806 Advanced Natural Language Processing
- 6.883 Online Machine Learning
- 6.849 Geometric Folding Algorithms

Graduate

- Introduction to Network Science
- Network Science II
- Social Networks
- Bayesian and Network Statistics

Publications

Papers in conference proceedings

1. Bhatia, V., V. P. Akavoor, S. Paik, L. Guo, M. Jalal, A. Smith, D. A. Tofu, E. E. Halim, Y. Sun, M. Betke, P. Ishwar, and D. T. Wijaya (Nov. 2021). OpenFraming: Open-sourced Tool for Computational Framing Analysis of Multilingual Data. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, pp.242–250. <https://aclanthology.org/2021.emnlp-demo.28> (visited on 05/13/2022).

Conference posters

1. Smith, Alyssa (Aug. 2020). *Doing Data Science With Queer Identities*. Poster. <https://cssh.northeastern.edu/nulab/pacss/2020-posters/>.

Workshops

1. Smith, A. (Aug. 2021). *infograffiti*. original-date: 2020-09-20T21:51:39Z. <https://github.com/lentilsoup/infograffiti> (visited on 05/13/2022).