

Laure Thompson

Research Software Engineer

Center of for Digital Humanities
Princeton University

Email: laurejt@princeton.edu

Website: <https://laurejt.github.io/>

Research Interests

Natural language processing, machine learning, digital humanities, computational social science

Education

PhD, Computer Science, Cornell University. 2020

Thesis: Understanding and Directing What Models Learn | Minor: Classical Archaeology

Committee: David Mimno (chair), David Bamman, Caitlín Barrett, Dexter Kozen

BS, Computer Science & BS, Electrical Engineering, University of Washington. 2013

Cum Laude. EE Concentration: Digital Signal Processing | Minors: Mathematics & Classical Studies

Positions

Academic

- **Research Software Engineer, Princeton University.** 2024-Present
Center for Digital Humanities
- **Adjunct Assistant Professor, University of Massachusetts Amherst.** 2024-Present
Manning College of Information and Computer Sciences
- **Assistant Professor, University of Massachusetts Amherst.** 2020-2024
Manning College of Information and Computer Sciences
- **Visiting Assistant Professor, University of Berkeley.** Fall 2023
School of Information

Industry

- **Microsoft Research: Research Intern.** Summer 2016
Systems Research Group – Redmond, WA
- **Google: Software Engineer Intern.** Summer 2012
DoubleClick Bid Manager – Kirkland, WA
- **Microsoft: Software Development Engineer in Test Intern.** Summer 2011
Windows Engineering Systems and Compatibility – Redmond, WA

Publications

Journal Articles

- **Computational Cut-Ups: The Influence of Dada.** [Laure Thompson](#) and David Mimno. *Journal of Modern Periodical Studies*. The Journal of Modern Periodical Studies Vol.8, No.2, 2018, pp. 179-195.

Peer-reviewed Conference Papers

- **Latin Treebanks in Review: An Evaluation of Morphological Tagging Across Time.** Maria Hudspeth, Brendan O'Connor, [Laure Thompson](#). Workshop on Machine Learning for Ancient Languages (2024) ML4AL.
- **Modifying Exemplification in Long-form Question Answering via Retrieval.** Shufan Wang, Fangyuan Xu, [Laure Thompson](#), Eunsol Choi, and Mohit Iyyer. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL), 2022.
- **Phrase-BERT: Improved Phrase Embeddings from BERT with an Application to Corpus Exploration.** Shufan Wang, [Laure Thompson](#), and Mohit Iyyer. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2021.
- **Authorless Topic Models: Biasing Models Away from Known Structure.** [Laure Thompson](#) and David Mimno. International Conference on Computational Linguistics (COLING), 2018. *Best Paper Award: Best NLP Engineering Experiment*.
- **The Strange Geometry of Skip-Gram with Negative Sampling.** David Mimno and [Laure Thompson](#). Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017. *Outstanding Paper (Best Paper Honorable Mention)*.
- **Quantifying the Effects of Text Duplication on Semantic Models.** Alexandra Schofield, [Laure Thompson](#), and David Mimno. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017.
- **Vale: Verifying High-Performance Cryptographic Assembly Code.** Barry Bond, Chris Hawblitzel, Manos Kapritsos, Rustan Leino, Jay R. Lorch, Bryan Parno, Ashay Rane*, Srinath Setty, and [Laure Thompson](#). USENIX Security Symposium, 2017. *Distinguished Paper Award*.
- **A Coalgebraic Decision Procedure for NetKAT.** Nate Foster, Dexter Kozen, Mae Milano, Alexandra Silva, and [Laure Thompson](#). ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL), 2015.

Book Chapters

- **Humanities and Human-Centered Machine Learning.** [Laure Thompson](#) and David Mimno. In *Human-Centered Machine Learning*, ed. Rebecca Fiebrink, Marco Gillies, Gonzalo Ramos. Forthcoming.

Other

- **Topic Modeling with Contextualized Word Representation Clusters.** [Laure Thompson](#) and David Mimno. Preprint: arXiv:2010.12626

- **Progressive Fusion for Multimodal Integration.** Shiv Shankar, [Laure Thompson](#), and Madalina Fiterau. Preprint: arXiv:2209.00302

Talks

Invited Talks and Panels

- **Computational Humanities & Human-Centered Machine Learning.** Computational Social Science Institute (CSSI) & Computational Humanities Initiative (CHI) joint seminar, University of Massachusetts Amherst, November 2022.
- **Computing for the Common Good.** Invited Panelist, Interdisciplinary Conversation, University of Massachusetts Amherst, February 2022.
- **Magic and Its Limits: Computational Categorization and Magical Gems.** The Future of the Past: Classics + Technology Workshop Series, Stanford Humanities Center, November 2021.
- **Studying Science Fiction at Scale.** Workshop on Narrative Understanding (WNU), June 2021.
- **A Symbiotic Future for Machine Learning and the Humanities.** Center for Humanistic Inquiry (CHI) Salon, Amherst College, March 2021.
- **Data Science for the Common Good.** Invited Panelist, Voices of Data Science at UMass Amherst, February 2021.
- **Interdependence in Action: The “Advanced Collaborative Support” Program from the HathiTrust Research Center.** Invited Panelist. HathiTrust Community Week, October 2020.
- **Understanding and Directing What Models Learn.** Machine Learning and Friends Lunch seminar, University of Massachusetts Amherst, October 2019.
- **Understanding and Directing the Learning of Latent Vector Spaces.** Graduate student colloquium, Rutgers University, April 2019.
- **Data Analysis for the Humanities.** Cornell Summer Graduate Fellowship in Digital Humanities program, June 2018.
- **Bilingual Topic Modeling in the *Patrologia Graeca*.** With David Mimno. Future Philologies: Digital Directions in Ancient World Text, April 2018.

Conference and Workshop Presentations

- **Reinventions and Responsibilities in the Age of AI.** Moderator. Digital Humanities (DH) Conference, August 2024.
- **RSE Turn in Digital Humanities.** Panelist. DARIAH Annual Event, June 2024.
- **The Life Cycles of Text Data: A Discussion on Access, Circulation, & Control.** Moderator. Text as Data (TADA), November 2023.
- **Text as Data.** Discussant, Text as Data (TADA), October 2022.
- **What Do Contextualized Representations Represent?** Text as Data (TADA), October 2021.
- **Finding Speculative Fiction in HathiTrust.** Association for Computers and the Humanities (ACH) Conference, July 2021.

- **Constructing and Analyzing Short Science Fiction at Scale.** Digital Humanities (DH) Conference, July 2020.
- **Analysis of Visual Corpora with Deep Learning.** Chair and panelist. Association for Computers and the Humanities (ACH) Conference, July 2019.
- **Medieval Mallet Mishaps: Topic Modeling Difficult Corpora.** With David Mimno and Anna Fore Waymack. 54th International Congress on Medieval Studies, May 2019.
- **Authorless Topic Models: Biasing Models Away from Known Structure.** Poster presentation. Ninth Annual Conference on New Directions in Analyzing Text as Data (TADA), September 2018.
- **Measuring Oracular Authenticity: Distinguishing the Historical from the Legendary in the Oracles of Delphi.** LAWNY Digital Antiquity Workshop at ISAW, December 2016.

Awards & Honors

- Selected Attendee for Rising Stars Workshop, 2019
- HathiTrust Research Center Advanced Collaborative Support with David Mimno, 2019
- Outstanding Teaching Assistant Award, 2019
- National Science Foundation Graduate Research Fellowship, 2014
- Cornell University Fellowship, 2013
- Gary Kildall Endowed Scholarship, 2012
- NASA Space Grant Scholarship, 2008-2012

Teaching Experience

Instructor.

University of Massachusetts Amherst – Amherst, NY

- Introduction to Social & Cultural Analytics: Spring 2021, Spring 2022
- Applications of Natural Language Processing: Fall 2021, Fall 2022, Fall 2023

Teaching Assistant.

Cornell University – Ithaca, NY

- Text Mining for History and Literature: Fall 2017
- System Security: Spring 2015, Spring 2019, Spring 2020

University of Washington – Seattle, WA

- Introduction to Compiler Construction: Winter 2013
- Software Design & Implementation: Winter 2012

Service

Conference Organization.

- Program Chair, Computational Humanities Research (CHR), 2024.
- Co-organizer, Text as Data (TADA), 2023

Program Committees and Reviewing.

Grant Reviewing

- Panelist, National Endowment for the Humanities, DHAG. 2021

Natural Language Processing Conferences

- Area Chair: ACL Rolling Review (2024), EMNLP (2022)
- Reviewer: ACL Rolling Review (2021, 2023), ACL (2021), EACL (2021)

Digital Humanities & Computational Social Sciences Conferences

- ACH (2021), CHR (2023), DH (2020, 2022-2024), DH Unbound (2022), TADA (2022-2023)

Workshops

- Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL; 2021-2022, 2024); Workshop on Narrative Understanding, Storylines, and Events (NUSE; 2020)

Ad-hoc Reviewing

- Journal of Cultural Analytics (CA; 2024)
- Sociological Methods and Research (SMR; 2023)

Institutional.

University of Massachusetts Amherst, College of Information & Computer Sciences

- PhD Admissions Committee Member. 2020-2021, Fall 2023
- Executive & Budget Committee Member. 2022-2023
- Teaching Faculty Search Committee Member. 2021-2022
- AFR Committee Member. 2021-2022
- CARE Mentoring Committee Member. 2020-2021

Cornell University, Computer Science

- Research Night Organizer. Spring and Fall 2019
- PhD Admissions Committee Member. 2018-2019
- Vice President, Computer Science Graduate Organization. Fall 2015
- Prospective Grad Student Visit Day Student Organizer. 2014
- Fellowship Archive and Workshop Organizer. 2013-2015

University of Washington, Computer Science & Engineering

- Event Coordinator, UW ACM Student Chapter. 2012-2013
- Public Relations Officer, UW ACM-W Student Chapter. 2011-2012

Community.

- Tech Savvy Workshop Leader. 2017
- Expanding Your Horizons Workshop Leader. 2015, 2016, 2017, 2018