Maria Antoniak

✉ maa343@cornell.edu
🌐 http://maria-antoniak.github.io

Research Interests

natural language processing • computational social science • cultural analytics • unsupervised methods • online communities • healthcare support • affect & narrative

Education

2016-Present **Ph.D. Information Science**, Cornell University
- Advisor: David Mimno
- Committee Members: Lillian Lee, Jeff Rzeszotarski, Richard Jean So
- Minor: Computer Science

2013-2014 **M.S. Computational Linguistics**, University of Washington
- Advisor: Fei Xia
- Thesis: *Extracting Topically Related Synonyms from Twitter*

2007-2011 **B.A. Program of Liberal Studies**, University of Notre Dame
- Glynn Family Honors Program

Publications

Tags, Borders, and Catalogs: Social Re-Working of Genre on LibraryThing
Maria Antoniak, Melanie Walsh, and David Mimno
Proceedings of the ACM on Human-Computer Interaction (CSCW), 2021

The Goodreads "Classics": A Computational Study of Readers, Amazon, and Crowd-sourced Literary Criticism
Melanie Walsh and Maria Antoniak
Post45 x Journal of Cultural Analytics, 2021

Imagined examples of painful experiences provided by chronic low back pain patients and attributed a pain numerical rating score
Robert Stewart Griffin, Maria Antoniak, Phuong Dinh Mac, Vladimir Kramskiy, Seth Waldman, David Mimno
Frontiers in Neuroscience, 2020

Narrative Paths and Negotiation of Power in Birth Stories
Maria Antoniak, David Mimno, and Karen Levy
Proceedings of the ACM on Human-Computer Interaction (CSCW), 2019

Evaluating the Stability of Embedding-based Word Similarities.
Maria Antoniak and David Mimno
Transactions of the Association for Computational Linguistics (TACL), 2018

Natural Language Processing Techniques on Oil and Gas Drilling Data
Maria Antoniak, Jeff Dagliesh, and Justin Lo
SPE Intelligent Energy Conference and Exhibition, 2016

Leveraging Paraphrase Labels to Extract Synonyms from Twitter
Maria Antoniak, Eric Bell, and Fei Xia
Proceedings of the Florida Artificial Intelligence Research Society Conference (FLAIRS), 2016

Workshop Presentations

Diverging Paths in Birth Stories: A Medical Dataset for Narrative Analysis.
Maria Antoniak and David Mimno.
Oral presentation at *Workshop for Narrative Understanding, NAACL*, 2019.

Comparative Analysis of Cultural Biases with Word Embeddings.
Maria Antoniak and David Mimno.
Poster presented at *Text as Data (TADA)*, 2018.

Evaluating the Stability of Embedding-based Word Similarities.
Maria Antoniak and David Mimno.
Poster presented at *Amazon Graduate Research Symposium*, 2017.

Extracting Topically Related Synonyms from Twitter.
Maria Antoniak, Eric Bell, and Fei Xia.
Poster presented at *Workshop for Women in Machine Learning (WiML)*, collocated with *NeurIPS*, 2015.

Panels

Replication and Computational Literary Studies.
Christof Schoch, Karina van Dalen-Oskam, Maria Antoniak, Fotis Jannidis, David Mimno.
Digital Humanities, 2020.

Cultural Analytics and the Book Review: Models, Methods, and Corpora.
Matthew J. Lavin, Kent Chang, Dan Sinykin, Melanie Walsh, Maria Antoniak, Kent Chang, Yuerong Hu, Wenyi Shang, Aniruddha Sharma, Shubhangi Singhal, Ted Underwood, Jessica Witte, and Peizhen Wu.
Digital Humanities, 2020.

Industry Experience

2020 Summer - 2021 Spring
Microsoft Research, Research Intern
- Mentors: Fernando Diaz, Asia Biega, Alexandra Olteanu
- Team: Fairness, Accountability, Transparency, Ethics (FATE)
- Self-disclosure and trust in online gaming communities.

2019 Summer
Facebook, Core Data Science Intern
- Mentors: Amit Bahl and Emma Ideal
- Unsupervised models of conversational roles in online health support and hiring groups.

2018 Summer
Microsoft Research, Research Intern
- Mentor: Ranjani Ramamurthy
- Unsupervised extractive summarization of clinical dialogues for Project EmpowerMD.

2014-2016
MAANA, Data Scientist
- Led the data science design and implementation of successful proof-of-concept projects for Fortune 100 companies (e.g., Shell, General Electric, Chevron).
- Implemented the NLP sections of a Scala machine learning pipeline.
- Created a custom named entity recognition/normalization system for unstructured text and a query expansion system for the core product.

2014 Summer
Pacific Northwest National Laboratory, National Security Intern
- Mentor: Eric Bell
- Twitter synonym extraction to improve paraphrase detection performance using machine translation metrics.
- Supported other NLP projects (sentiment analysis, topic modeling).

2013 Summer
Pacific Northwest National Laboratory, Mobile Developer Intern
- Mentor: Courtney Corley
- Created *FoodFeed*, an interactive foodborne illness application using public health data from the FDA and CDC, which won PNNL's biosurveillance mobile app competition.

Teaching Experience

Fall 2020	TA, **CS 3350: Text Mining for History and Literature**, Cornell University
Fall 2019	TA, **CS 3350: Text Mining for History and Literature**, Cornell University
Spring 2019	TA, **CS 3300: Data-Driven Web Applications**, Cornell University
Spring 2017	TA, **CS 3300: Data-Driven Web Applications**, Cornell University
2011-2012	English Teacher, **Ukrainian Catholic University**, Lviv, Ukraine
2010 & 2011	TA, **Humanities Spring**, Assisi, Italy
2010-2011	TA, **First Year Writing Tutorial**, University of Notre Dame
2008-2011	Writing Tutor, **University Writing Center**, University of Notre Dame

Honors, Awards, & Scholarships

2018	Information Science Outstanding Service Award, *Cornell University*
2016-2017	Information Science Fellowship, *Cornell University*
2013	1st Place, *PNNL Biosurveillance Mobile App Development Competition*
2007-2011	Glynn Family Honors Program, *University of Notre Dame*
2007	National Merit Scholar, *National Merit Scholarship Program*

Professional Service

Reviewed for *NAACL 2021*
Reviewed for *EACL 2021*
Reviewed for *CHI 2021*
Reviewed for *EMNLP 2020*

Mentoring

2019-2020	*LeAnn McDowall*, undergraduate, data analysis of birth control websites

Community Service

2020-2021	*President*, Graduate Students for Gender Inclusion in Computing
2020-2021	*Student Representative*, Broadening Participation in Information Science at Cornell
2020-2021	*Webmaster*, Cornell Graduate Women in Science
2019-2020	*President*, Graduate Students for Gender Inclusion in Computing
2019	*Lead Organizer*, Cornell Information Science Research Retreat
2019-2020	*Treasurer*, Cornell Graduate Women in Science
2018-2019	*Social Chair*, Cornell Graduate Women in Science
2018-2019	*Treasurer*, Information Science Graduate Student Association
2017-2018	*Vice President & Treasurer*, Information Science Graduate Student Association
2010-2011	*Editorial Board Member*, Notre Dame Journal of Undergraduate Research

Non-Degree Studies

2016 Summer	Lisbon Machine Learning Summer School (LxMLS), *Lisbon, Portugal*
2011 Summer	Ukrainian Language Summer School, *Lviv, Ukraine*
2010 Spring	French Language & Culture, Université Catholique de l'Ouest, *Angers, France*