Maria Antoniak

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

Research Interests

Broad   Natural language processing, machine learning

Narrow   Unsupervised methods (word embeddings, topic modeling), narrative structure, sentiment analysis, computational social science, medicine and health, digital humanities

Education

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

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

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

Publications

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.

Industry Experience

2019 Summer **FACEBOOK**
Core Data Science Intern
- Mentors: Amit Bahl and Emma Ideal
- Unsupervised modeling of conversational roles in online health groups and hiring groups.

2018 Summer **MICROSOFT**
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).
- Created custom NLP features and pipelines 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.

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 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*

Service

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*

Skills

Computing

Python Stack: spaCy, Gensim, scikit-learn, pandas, SciPy, NumPy, PySpark, matplotlib
Other: Mallet, Stanford CoreNLP
Limited/Older Experience: Java, C++, C, Scala, Javascript

Languages

English: native speaker
French: intermediate
Ukrainian & Italian: elementary