Maria Antoniak

✉ maa343@cornell.edu

🌐 http://maria-antoniak.github.io

Research Areas

Natural Language Processing • Computational Social Science • Cultural Analytics

Education

2016-Present **Ph.D. Information Science**, Cornell University
- Committee: David Mimno, 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

2021 Bad Seeds: Evaluating Lexical Methods for Bias Measurement
Maria Antoniak and David Mimno
Proceedings of the Annual Meeting of the Association for Computation Linguistics (ACL)

2021 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 Crowdsourced Amateur Criticism
Melanie Walsh and **Maria Antoniak**
Post45 x Journal of Cultural Analytics

2020 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

2019 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)

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

2016   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)

Industry Experience

2021   **Twitter**, *Cortex Intern*
       - Mentors: Stephen Ragain and Jay Baxter
       - Machine learning to support crowdsourced misinformation identification (Birdwatch)

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

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

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

2014-2016   **MAANA**, *Data Scientist*
       - Data science design and implementation of proof-of-concept projects
       - Product query, named entity recognition, and machine learning pipeline features

2014   **Pacific Northwest National Laboratory**, *National Security Intern*
       - Mentor: Eric Bell
       - Synonym extraction for paraphrase detection on Twitter
       - Supported other NLP projects (sentiment analysis, topic modeling)

2013   **Pacific Northwest National Laboratory**, *Mobile Developer Intern*
       - Mentor: Courtney Corley
       - Public health, foodborne illness, mobile application development

Workshop Presentations

2021   **Bad Seeds: Evaluating Lexical Methods for Bias Measurement**
       Maria Antoniak and David Mimno
       Poster at *Gender Bias Workshop*, collocated with *ACL*

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

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

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

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

Talks

2021   **BERT for Humanists: Uses, Challenges, Translations**
Maria Antoniak, Melanie Walsh, David Mimno, and Matthew Wilkens
Association for Computers and the Humanities (ACH)

2021   **Building Social Bonds in Online Gaming Communities**
Maria Antoniak and Alexandra Olteanu
Microsoft AI & Gaming Research Summit

2020   **Narrative Paths and Negotiation of Power in Birth Stories**
University of Michigan, Data Science/Computational Social Science Seminar

Panels

2021   **Computational Approaches to Book Reviews**
Matthew Lavin, Melanie Walsh, Maria Antoniak, Dan Sinykin, Jordan Pruett
Association for Computers and the Humanities (ACH)

2020   **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

Teaching Experience

Fall 2021   **INFO 2950: Introduction to Data Science**, Cornell University, TA
June 2021   **Computational Text Analysis**, Barnard College, Guest Lecture
Fall 2020   **INFO 3350: Text Mining for History and Literature**, Cornell University, TA
Fall 2019   **INFO 3350: Text Mining for History and Literature**, Cornell University, TA
Spring 2019   **CS/INFO 3300: Data-Driven Web Applications**, Cornell University, TA
Spring 2017   **CS/INFO 3300: Data-Driven Web Applications**, Cornell University, TA
2011-2012   **Ukrainian Catholic University**, Lviv (Ukraine), English Teacher
2010 & 2011   **Humanities Spring**, Assisi (Italy), TA
2010-2011   **First Year Writing Tutorial**, University of Notre Dame, TA

2008-2011   **University Writing Center**, University of Notre Dame, Writing Tutor

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

2021   Reviewed for *EMNLP*
Reviewed for *CSCW*
Reviewed for Student Research Workshop, *ACL*
Reviewed for *NAACL*
Reviewed for *EACL*
Reviewed for *CHI*
2020   Reviewed for *EMNLP*

Mentoring

2019-2021   *LeAnn McDowall*, undergraduate; data analysis of birth control websites; won Best Paper Award at the *Black in AI Workshop* at *NeurIPS* 2020.

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

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