

# 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

### **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.  
*The Twenty-Eighth International 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 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, <b>CS 3350: TEXT MINING FOR HISTORY AND LITERATURE</b> , Cornell University
Spring 2019	TA, <b>CS 3300: DATA-DRIVEN WEB APPLICATIONS</b> , Cornell University
Spring 2017	TA, <b>CS 3300: DATA-DRIVEN WEB APPLICATIONS</b> , Cornell University
2011-2012	English Teacher, <b>UKRAINIAN CATHOLIC UNIVERSITY</b> , Lviv, Ukraine
2010 & 2011	TA, <b>HUMANITIES SPRING</b> , Assisi, Italy
2010-2011	TA, <b>FIRST YEAR WRITING TUTORIAL</b> , University of Notre Dame
2008-2011	Writing Tutor, <b>UNIVERSITY WRITING CENTER</b> , University of Notre Dame

## Honors, Awards, & Scholarships

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

## Service & Leadership

2019	Lead Organizer, <i>Cornell Information Science Research Retreat</i>
2019-2020	Treasurer, <i>Cornell Graduate Women in Science</i>
2018-2019	Social Chair, <i>Cornell Graduate Women in Science</i>
2018-2019	Treasurer, <i>Information Science Graduate Student Association</i>
2017-2018	Vice President & Treasurer, <i>Information Science Graduate Student Association</i>
2010-2011	Editorial Board Member, <i>Notre Dame Journal of Undergraduate Research</i>

## Non-Degree Studies

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

## 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

Last updated: November 10, 2019 •