MARIA ANTONIAK

maa343@cornell.edu

EDUCATION

Cornell University   *August 2016 - Present*
Ph.D. in Information Science
Adviser: David Mimno

Relevant Courses: Advanced Topic Modeling, Behavior and Information Technology

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

Relevant Courses: Artificial Intelligence, Grammar Engineering, Shallow Processing Techniques, Deep Processing Techniques, Advanced Statistical Methods for NLP

University of Notre Dame   *August 2007 - May 2011*
B.A. in Program of Liberal Studies
Glynn Family Honors Program
Adviser: Bernd Goehring

RESEARCH INTERESTS

Natural language processing, machine learning, unsupervised and semi-supervised learning, word embeddings, topic modeling, synonym extraction, paraphrase detection, digital humanities.

INDUSTRY EXPERIENCE

MAANA   December 2014 - May 2016
Data Scientist   *Seattle, WA*

· Led the data science design and implementation of three successful proof-of-concept projects for Fortune 100 companies, using various NLP and ML techniques (e.g. semi-supervised classification, topic modeling).
· Created a custom named entity recognition and normalization system for messy, technical text using bootstrapping and word embeddings.
· Constructed a query expansion and completion search system for the core product.
· Implemented the NLP sections of a Scala machine learning pipeline for the core product.

Pacific Northweset National Laboratory   July - December 2014
National Security Intern   *Seattle, WA*

· Researched automatic synonym extraction for Twitter to boost performance for paraphrase detection.
· Implemented two baseline systems for paraphrase detection using SVM models trained using machine translation metrics, a re-weighting scheme (TF-KLD), and non-negative matrix factorization (NMF).
· Supported other NLP software development projects, including sentiment analysis, topic modeling, and data extraction.

Pacific Northwest National Laboratory   June - Auguest 2013
Mobile App Developer Intern   *Richland, WA*

· Designed, created, and presented FoodFeed, a foodborne illness tracking application that won PNNL's biosurveillance mobile app competition.
· Developed the front-end of the Android application for Nexus 4 and 7.

· Scraped and indexed public health data from FDA and CDC using Solr/Lucene.

SERVICE AND TEACHING EXPERIENCE

Ukrainian Catholic University   Auguest 2011 - July 2012
Volunteer English Teacher   *Lviv, Ukraine*

· Taught English to university students and seminarians.
· Led the university's English club and organized activities for 100+ students.

Humanities Spring   Summers 2010 and 2011
Teaching Assistant   *Assisi, Italy*

· Taught a poetry and creative writing course for high school students.
· Assisted courses in Latin, Greek, art history, opera, and classic literature.
· Led student excursions to historical sites in Umbria and Tuscany.

First Year Writing Tutorial, University of Notre Dame   September 2010 - May 2011
Teaching Assistant   *Notre Dame, IN*

· Taught university-level writing skills to a core group of struggling freshmen.
· Conducted both individual sessions (6-8 per week) and either assisted or led weekly group sessions.
· Worked one-on-one with high profile athletes (ND Football) and ESL students.

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

· Directed 70+ discussion-based consultations per semester with individual students.
· Led specialized workshops for grant applicants.
· Completed one semester training course, with continued monthly training led by experts in various writing styles and disciplines (e.g. chemistry, philosophy, ESL).

PUBLICATIONS

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

- Antoniak, Maria, Eric Bell, and Fei Xia. "Leveraging Paraphrase Labels to Extract Synonyms from Twitter." *The Twenty-Eighth International Florida Artificial Intelligence Research Society Conference.* 2015.

POSTERS

- Antoniak, Maria, Eric Bell, and Fei Xia. "Extracting Topically Related Synonyms from Twitter." In *Workshop for Women in Machine Learning (WiML),* collocated with NIPS 2015.

TECHNICAL SKILLS

Computer Languages	Python, Java, C++, Scala
Python Stack	spaCy, Gensim, scikit-learn, pandas, SciPy, NumPy, PySpark
Other Toolkits	Mallet, Stanford CoreNLP, NLTK, openNLP, Solr/Lucene
Random	SVN, Git
Human Languages	English (native), French (intermediate), Ukrainian (beginning)
	Italian (beginning), Russian (beginning)

HONORS AND AWARDS

- **Information Science Fellowship**
  Cornell University, 2016
  Awarded to applicants who show particular ability, promise, and fit to the IS program.

- **1st Place, Biosurveillance Mobile App Development Competition**
  Pacific Northwest National Laboratory, 2013
  Designed and implemented an award winning mobile app to increase public food safety.

- **Editorial Board Member, Journal of Undergraduate Research**
  University of Notre Dame, 2010
  Selected by peers and deans.

- **Glynn Family Honors Program**
  University of Notre Dame, 2007 - 2011
  Top 5% of freshman with talent in both the arts and sciences.

- **National Merit Scholar**
  National Mertit Scholarship Program, 2007
  Top 0.5% of scores on the PSAT.

OTHER ACADEMIC EXPERIENCE

- **Lisbon Machine Learning Summer School (LxMLS)**
  Instituto Superior Tecnico (Lisboa), July 2016