Tom Effland

Contact teffland@cs.columbia.edu teffland.github.io teffland tomeffland

Research Interests

Natural language processing, machine learning, semi-supervised and active learning, grounded knowledge discovery, intelligent assistance, decision support systems.

Education

Columbia University, New York, New York USA

Ph.D., Computer Science, Expected May, 2020
M.S., Computer Science, December, 2016
Research Area: Information Extraction, Natural Language Processing, Machine Learning
Advisers: Prof. Michael Collins and Prof. Daniel Hsu

University at Buffalo, The State University of New York, Buffalo, New York USA

B.S., Applied Mathematics, Honors, May, 2015
Minors in Computer Science and Statistics, *(GPA: 3.99/4.0)*

Research Experience

Columbia University, New York, New York USA

Graduate Research Assistant **September, 2017 - Present**
Department of Computer Science
- Researching semi-supervised learning of broad-coverage grounded semantic parsing using variational bayesian methods, with a focus on leveraging existing structured knowledge
- *Project Advisers:* Prof. Michael Collins and Prof. David Blei

Graduate Research Assistant **August, 2015 - July, 2017**
Department of Computer Science
- Researched principled methods and strategies for extraction of actionable information from rare events on social media
- Collaborated with NYC Department of Health to identify foodborne illness outbreaks from Yelp and Twitter to facilitate targeted investigation of restaurants
- *Project Advisers:* Prof. Luis Gravano and Prof. Daniel Hsu

TextIQ, Inc., New York, New York USA

Research Intern **June, 2016 - August, 2016**
- Researched novel techniques for natural language interfaces, semantic parsing, and question answering over structured knowledge sources
- Implemented a system for automatically building question answering chatbots from scratch for new structured data sources

University of Illinois at Urbana-Champaign, Urbana, Illinois USA

Research Assistant **June, 2014 - August, 2014**
Passionate on Parallel NSF-Supported REU, Parallel Computing Institute
- Researched techniques for using Hadoop to automatically parallelize scientific codes
- Parallelized atmospheric science pollution simulation research software with MPI
- *Project Advisers:* Prof. Nicole Riemer, Prof. Matthew West

University at Buffalo, The State University of New York, Buffalo, New York USA

Independent Honors Research **August, 2014 - April, 2015**
Department of Computer Science
- Won 1st place in ACM SIGCSE Undergraduate Student Research Competition Grand Finals

- Researched and developed context-focused web crawling framework for extracting similar content from heterogeneous seed domains.
- Specific application was retrieval of university course descriptions given only domain names.
- *Project Adviser:* Prof. Bina Ramamurthy

NASA Europa Challenge Team Member, iGlobe project **March, 2014 - May, 2014**
Department of Computer Science
- Won 2nd place University Project in international software competition
- Researched and coordinated implementation of weather API interface layer into iGlobe
- *Project Adviser:* Prof. Varun Chandola

Research Assistant **January, 2013 - May, 2014**
Department of Computer Science, Department of Mathematics
URGE to Compute NSF-Supported REU
- Developed scalable, accurate, and secure matching algorithms for fingerprints
- Researched machine learning and error correcting code applications to secure fingerprint matching
- *Project Advisers:* Prof. Atri Rudra, Prof. John Ringland

Professional Experience

Instabeat, Inc., San Francisco, California, USA

Technical Consultant **June, 2018 - September 2018**
- Designed, developed, and deployed machine learning pipeline for supporting app technology

NYC Department of Health and Mental Hygiene, New York, New York, USA

Technical Consultant **January, 2016 - January, 2018**
- Designed and developed social media analysis machine learning pipeline for automatically detecting possible foodborne illness outbreaks, supporting epidemiologist investigations

Schussmeisters Ski Club, Inc., Buffalo, New York, USA

Technical Consultant **May, 2013 - April, 2015**
- Synthesized business needs and developed integrated data management application to provide business organization and analytics for informing data-driven decisions by board members

Journal Publications

T. Effland, Anna Lawson, Sharon Balter, Katelynn Devinney, Vasuhda Reddy, Luis Gravano, Daniel Hsu. Discovering Foodborne Illness in Online Restaurant Reviews. *Journal of the American Medical Informatics Association, Volume 25, Issue 12, 1 December 2018, Pages 1586 - 1592, https://doi.org/10.1093/jamia/ocx093*

Conference Publications

T. Effland. 2015. Focused Retrieval of University Course Descriptions from Highly Variable Sources. In *ACM Student Research Competition Undergraduate Grand Finals*. **First Place Award**.

J. Hartloff, M. Morse, B. Zhang, **T. Effland**, J. Cordaro, J. Schuler, S. Tulyakov, A. Rudra, V. Govindaraju. 2015. A Multiple Server Scheme for Fingerprint Fuzzy Vaults. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*.

M. Morse, J. Hartloff, **T. Effland**, J. Schuler, J. Cordaro, S. Tulyakov, A. Rudra, V. Govindaraju. 2014. Secure Fingerprint Matching With Generic Local Structures. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on*. pages 84-89.

T. Effland, M. Schneggenburger, J. Schuler, B. Zhang, J. Hartloff, J. Dobler, S. Tulyakov, A. Rudra, V. Govindaraju. 2014. Secure fingerprint hashes using subsets of local structures. In *Proc. SPIE 9075-12, Biometric and Surveillance Technology for Human and Activity Identification XI, 90750D*.

Conference Presentations

T. Effland. Focused Retrieval of University Course Descriptions from Highly Variable Sources. ACM SIGCSE Student Research Competition, Kansas City, Missouri, March, 2015.

T. Effland, M. Schneggenburger, J. Schuler, B. Zhang, J. Hartloff, J. Dobler, S. Tulyakov, A. Rudra, V. Govindaraju. Secure fingerprint hashes using subsets of local structures. SPIE Defense, Sensing, Security Biometrics Workshop, Baltimore, Maryland, May, 2014.

T. Effland, M. Schneggenburger, J. Schuler. Fingerprints as Passwords. National Conference for Undergraduate Research (NCUR), Lexington, Kentucky, April, 2014.

Poster Presentations

Drashko Nakikj, **T. Effland**. The Posts Recommendation Algorithm for dExplorer. Columbia University Data Visualization and Exploration Poster Event, New York, New York, May, 2017.

T. Effland. Identifying Foodborne Illness from Social Media. Columbia University Data Science Day, New York, New York, April, 2016.

T. Effland. Focused Retrieval of University Course Descriptions from Highly Variable Sources. University at Buffalo Celebration of Excellence, Buffalo, New York, April, 2015.

T. Effland. Focused Retrieval of University Course Descriptions from Highly Variable Sources. ACM SIGCSE Student Research Competition, Kansas City, Missouri, March, 2015.

Technical Skills

Languages: Python , Javascript
Libraries & Frameworks: PyTorch, Tensorflow, Pandas, SpaCy, Jupyter, Scikit-Learn
Web & Visualization: React, MatPlotLib, HTML, CSS, jQuery, d3
Data: JSON, XML, SQL, MongoDB, PostgreSQL, Neo4J, ElasticSearch

Honors and Awards

Columbia University:

- Northeast Big Data Hub Young Innovators Award **June, 2016**
- NSF Graduate Research Fellowship **April, 2016**
- NSF IGERT "From Data to Solutions" Fellowship **August, 2015**

University at Buffalo:

- 1st Place - ACM Student Research Competition Grand Finals **May, 2015**
- Outstanding Senior Award, Mathematics Department **May, 2015**
- NSF Data-Intensive Computing Fellowship (NSF-DUE-CCLI-0920335) **August, 2014**
- 2nd Place - NASA Europa International Software Competition **June, 2014**
- Harriet F. Montague Award **May, 2014**
- Phi Beta Kappa **February, 2014**
- Grace Capen Academic Award **May, 2013**
- Provost Scholarship **Fall 2011 - Spring 2015**
- Deans List **Fall 2011 - Spring 2015**