Justin Sybrandt

Ph.D. Candidate

Email: Justin@Sybrandt.com Website: Justin.Sybrandt.com
GitHub: JSybrandt Phone: (484) 354-8692

Objective Putting unstructured information to good use remains a challenge, especially regarding scientific text. With every day comes thousands of newly published papers, which can lead to missed findings and, in turn, delayed progress. In order to aid researchers, my work involves understanding this information through text-based and graph-based machine learning methods in my thesis project: Moliere.

Education

Clemson University (Aug. 2016 - May 2020)

- Ph.D. in computer science. (GPA 4/4)
- Relevant Coursework: Design & Analysis of Algorithms, Advanced Data Structures, Deep Learning, Data Mining, Distributed & Cluster Computing, Parallel Architecture, Network Science
- Recipient of the GAANN DAISE & NRT PhD. fellowship.
- Member of the ACM and IEEE

Grove City College (Aug. 2012 - May 2016)

- BS in computer science, minor in mathematics.
- Graduated Summa Cum Laude (GPA 3.85/4).
- Top of class in computer science (in-major GPA 3.95/4).

Research Interests

- Machine Learning
- Text Mining
- Hypothesis Generation
- Natural Language Understanding
- Graph Mining
- Artificial Intelligence

Work Experience

Summer 2019, Software Engineer Intern, Facebook

- Accepted as a Ph.D. Intern at Instagram in New York City.

Summer 2018, Ph.D. SWE Intern, Google

- Designed a graph-mining solution for identifying product attributes.
- Implemented solution at scale via Google-internal distributed systems.
- Performed comprehensive validation, ensuring classifier performance across product categories.

Summer 2017, Graduate Research Assistant - Los Alamos National Lab

- Developed high performance software in julia for non-negative matrix factorization.
- Extended MOLIERE to water resources research with the computational environmental science group (EES-16).

2015-2016, Programming Intern - Vigilant Cyber Systems, Inc.

- Developed a visualization library in Scala using ScalaFX.
- Independently managed time when working remotely.
- Balanced senior-level course work with development.

Summer 2015, Student Researcher - UC Berkeley & NERSC

- Designed and implemented a tool to quickly synchronize multi-petabyte General Parallel File Systems.
- Presented a poster at the ACM Student Poster Session at SC'15.
- Presented a work in progress paper at the Parallel Data Storage Workshop.

Summer 2014, Student Researcher - Grove City College

- Added distribution preferences to Data Stream Management Systems (DSMS).

- Simulated new DSMS features in Python.
- Studied modern DSMS through recent research papers.
- Presented a poster at the Grove City student poster session.

2012-2014, Programming Intern - Gravic Inc.

- Worked on a six person team developing tools for administering exams.
- Collaborated with corporate partners to design features which allowed our products to share data.
- Gained familiarity with project management while extending the Remark VB.NET code base.

Development Skills and Technologies

- C++
- SQL
- Linux
- Python
- MongoDB
- VIM
- Keras
- Scala
- LaTeX
- Bash
- Julia
- Git

Publications

Peer-Reviewed Papers

- Sybrandt, J., Shtutman, M., & Safro, I. (2017, August). Moliere: Automatic biomedical hypothesis generation system. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1633-1642). ACM.
  - Acceptance rate 8.8%
  - Audience appreciation award honorable mention (in top 5)
  - Awarded best project at Clemson's data science & network analysis student conference by Drs. B. Dean, D. House, and L. Ngo.
- Sybrandt, J., Shtutman, M., & Safro, I. (2018, December). Large-scale validation of hypothesis generation systems via candidate ranking. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 1494-1503). IEEE.
  - Acceptance rate 18%
- Sybrandt, J., Carrabba, A., Herzog, A., & Safro, I. (2018, December). Are abstracts enough for hypothesis generation? In 2018 IEEE International Conference on Big Data (Big Data) (pp. 1504-1513). IEEE.
  - Acceptance rate 18%

Online Preprints & In-Submission Works

- Sybrandt, J., & Safro, I. A tale of two embeddings: Embed hypergraphs with different member and community structure. In-submission and not avalible online.
- Locke, W., Sybrandt, J., Safro, I., & Atamturktur, S. (2018, November 12). Using Drive-by Health Monitoring to Detect Bridge Damage Considering Environmental and Operational Effects. https://doi.org/10.31224/osf.io/ntfdp
- Shaydulin, R., & Sybrandt, J. (2017). To Agile, or not to Agile: A Comparison of Software Development Methodologies. arXiv preprint arXiv:1704.07469.

Peer-Reviewed Extended Abstracts

- Aksenova, M., Sybrandt, J., Cui, B., Lucius, M., Ji, H., Wyatt, M., Safro, I., Zhu, J., & Shtutman, M. (2019). Inhibition of the DEAD Box RNA Helicase 3 prevents HIV-1 Tat- and cocaine-induced neurotoxicity by targeting microglai activation. In 2019 Meeting of the NIDA Genetic Consortium. Extended Abstract & Poster
- Sybrandt, J., & Hick, J. (2015). Rapid replication of multi-petabyte file systems. Work in progress in the 2015 Parallel Data Storage Workshop. Poster in 2015 Super Computing.

Teaching Experience

- Spring 2018 & Fall 2018, Guest Lecture: Applied Data Science
- Fall 2017 - Spring 2018, Project Supervisor for senior projects in Seminar in Professional Issues II

Honors and Awards

- Member of Upsilon Pi Epsilon CS Honor Society.
- Member of the Kappa Mu Epsilon National Mathematics Honor Society.
- Member of the Alpha Tau Mu chapter of Mortarboard, a service-oriented honor society.
- Recipient of the KDD'17 ACM student travel award.
- Recipient of the BigData'18 IEEE student travel award.
- Recipient of the Clemson CCIT Super Computing'18 travel award.