Justin Sybrandt

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

Objective   I am looking for cutting-edge research that impacts both the state-of-the-art as well as the everyday lives of people worldwide. I have a significant amount of research experience in text- and graph-embedding, as well as practical experience at Google and Facebook. My in-progress Ph.D. enabled me to propose new ML algorithms, as well as design the *Moliere* system, which learns to predict fruitful new research directions from previously published medical journals.

Education

Ph.D. in CS, Clemson University (Aug. 2016 - May 2020)

- Thesis: *Exploiting Latent Features of Text and Graphs*
- Relevant Coursework (GPA 4/4): Design & Analysis of **Algorithms**, Advanced **Data Structures**, Data Mining, Distributed & Cluster Computing, Parallel Architecture, Network Science.
- Recipient of the GAANN DAISE & NRT PhD. fellowships.

BS in CS, Minor in Math, Grove City College (Aug. 2012 - May 2016)

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

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%.
- 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. (2019). FOBE and HOBE: First-and High-Order **Bipartite Embeddings**. arXiv preprint arXiv:1905.10953.
- Sybrandt, J., Shaydulin R., & Safro I. (2019). **Partition Hypergraphs** with Embeddings. Not yet available.
- 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
- Aksenova M., Sybrandt J., Cui B., Sikirzhytski V., Ji H., Odhiambo D., Lucius M., Turner J. R., Broude E., Pena E., Lizzaraga S., Zhu J., Safro I., Wyatt M. D., Shtutman M. (2019). **Inhibition of the DDX3** prevents HIV-1 Tat and cocaine-induced neurotoxicity by targeting microglia activation. https://www.biorxiv.org/content/10.1101/591438v1
- Shaydulin, R., & Sybrandt, J. (2017). To Agile, or not to Agile: A Comparison of **Software Development** Methodologies. arXiv preprint arXiv:1704.07469.

Development Skills and Technologies

Programming Languages	C++, Python, Bash, SQL, Matlab, Java, Scala
Tools	Git, Linux, VIM, LaTeX, Mercurial
ML-Libraries	Tensorflow, Keras, Scikit-Learn
Parallel/Distributed Programming	OpenMP, Spark, Flume, Dataswarm, GNU-Parallel, MPI

Work Experience

Summer 2019, Ph.D. SWE, Facebook [Current]

- Improved the detection of violating media on Instagram by leveraging **embeddings**.
- Demonstrated high productivity and fast learning speed, as evidenced by formal peer feedback, while adapting to the workflow at Facebook.

Summer 2018, Ph.D. SWE Intern, Google

- Proposed and produced a **graph-mining** solution for identifying product attributes that could decreased the need for human oversight by over 50%.
- Worked efficiently, developing my proposed system from a whiteboard idea to an in-production pipeline ahead of schedule.
- Performed comprehensive validation, ensuring classifier performance across product categories.
- Presented work to senior research scientists within Google's graph-mining team.

Summer 2017, Graduate Research Assistant, Los Alamos National Lab

- Developed high performance software in **Julia** for non-negative matrix factorization to be released in the open-source scientific computing library *madsjulia*.
- Evaluated the ability for my research project MOLIERE to extend to water resources research with the computational environmental science group.

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

- Developed a visualization library in **Scala** using ScalaFX for a DoD contract.
- Worked independently as a self-led remote employee, while simultaneously finishing my last year of undergrad.

Summer 2015, Student Researcher, UC Berkeley & NERSC

- Designed a tool in **Java** to quickly synchronize **multi-petabyte Parallel File Systems**.
- Created a poster that was accepted at the ACM Student Poster Session at the Supercomputing conference in 2015.
- Presented a work in progress paper at the Parallel Data Storage Workshop.

Summer 2014, Student Researcher, Grove City College

- Prototyped multiple temporal distribution preferences for **Data Stream Management Systems**.
- Implemented research ideas using **Python**, and deployed large-scale simulations on Linux.
- Created a poster that was presented at the Grove City student poster session.

2012-2014, Programming Intern, Gravic Inc.

- Worked closely with a small team to create tools for teachers to more easily administer exams.
- Implemented new features using **VB.NET** to allow for **OCR** detection of student IDs, fractional responses, and free-form textual answers.
- Collaborated directly with corporate partners to implement features that allowed our products to interoperate.