

James A. Scharf

(914) 522-2930 | jscharf8@jhu.edu | <https://github.com/JamesScharf>

EDUCATION

The Johns Hopkins University – Baltimore, MD

Expected May 2022

- Master of Science in Engineering, Computer Science

The Johns Hopkins University – Baltimore, MD

May 2021

- Bachelor of Arts, Political Science with minor in Computer Science
- Coursework in natural language processing, artificial intelligence, ethical hacking, and machine translation (deep learning)
- **GPA:** 3.56 (Awarded General Honors)

PUBLICATIONS AND PRESENTATIONS

James A. Scharf (presenter), Arya D. McCarthy and Giovanna Maria Dora Dore, “Evolution and Bias: News Portrayal of Civil Unrest in Hong Kong 1998-2020,” (Forthcoming in) *4th Politics and Computational Social Science Conference (PaCSS 2021)*.

James A. Scharf, Arya D. McCarthy and Giovanna Maria Dora Dore. “Characterizing News Portrayal of Civil Unrest in Hong Kong, 1998-2020.” (Forthcoming in) *Proceedings of the Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)*, August 2021. Bangkok, Thailand (Virtual).

David Yarowsky (presenter) et al. (including **James Scharf**), “A 1000-language Collaborative Universal Dictionary and Universal Translator,” UNESCO International Conference Language Technologies for All (LT4All).

EMPLOYMENT & RESEARCH EXPERIENCE

Joint MISO Web Operations Center (JMWC), U.S. Special Operations Command – Tampa, FL

Summer 2021

Civic Digital Fellow, Data Science

- Building automated fine-tuning multilingual named entity recognition and text classification models for self-designed CI/CD model pipeline with the Huggingface Transformers Library and SpaCy
- Utilizing Prodigy and Plotly’s Dash library for the creation of user interfaces

Department of East Asian Studies, Johns Hopkins University – Baltimore, MD

Summer 2020 – Present

Computational Research Assistant for Dr. Giovanna Maria Dora Dore

- Applied named entity recognition and sentiment analysis with Huggingface Transformers Library and topic modeling with Gensim to 4000 news articles (Forthcoming first author publication and primary presenter listed above)
- Generated plots with Seaborn and Matplotlib to compare news coverage in Hong Kong by year

Department of Computer Science, Johns Hopkins University – Baltimore, MD

Spring 2020 & Spring 2021

Head Course Assistant for Information Retrieval and Web Agents

- Assistant for a course in machine learning, information extraction and web scraping
- Helped students with projects in sense disambiguation, word embeddings and vector models of information retrieval
- Created and graded homework assignments

Center for Language and Speech Processing (CLSP), Johns Hopkins University – Baltimore, MD

Summer 2020

Dr. David Yarowsky’s Low-Resource Languages Lab ([LoReLab](#))

- Utilized Facebook’s Fairseq Sequence-to-Sequence (Seq2Seq) library for G2P tasks
- Scraped and parsed foreign language webpages into structured data with Pandas, Selenium and BeautifulSoup to help create the largest multilingual dataset of pronouns in the world in 1000+ languages (*publication pending submission*)

Department of Political Science, Johns Hopkins University – Baltimore, MD

September 2018 – October 2019

Research Assistant for Prof. Adam Sheingate

- Used Python and R to analyze F.E.C. data related to elections from 2010 to 2016
- Built a natural language classifier using the Scikit-learn machine learning framework, pattern matching and query-expansion

ODN (formerly Open Data Nation) – New York, NY

Summer 2019

Data Science Intern

- Cleaned datasets and performed statistical analyses of roadway and car crash data with Pandas and Scikit-learn
- Created maps and visualizations using Geopandas, Geoplot, Seaborn and Matplotlib