James "Jamie" Scharf

22 Garey Drive | Chappaqua, NY 10514 | Mobile: (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; Minor, Computer Science
- Coursework in natural language processing, artificial intelligence, and parallel programming
- **GPA:** 3.56 (General Honors)

Experience with the following programming languages: Python, C/C++, Java, Go, JavaScript

RECENT PUBLICATIONS & PRESENTATIONS

James A. Scharf, Arya D. McCarthy and Giovanna Maria Dora Dore. "Evolution and Bias: News Portrayal of Civil Unrest in Hong Kong, 1998-2020," *Politics and Computational Social Science (PaCSS) Conference*, August 2021. Boston, U.S.A. (Virtual).

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

Please see my Google Scholar profile for other publications.

EMPLOYMENT & RESEARCH EXPERIENCE

Joint MISO WebOps Center (JMWC), U.S. Special Operations Command – Tampa, FL   *Summer 2021*
Telesis Civic Digital Fellow, Data Science
- Created a multilingual in-house Named Entity Recognition (NER), Text Classification annotation platform for building and using customized neural and non-neural models in 400 languages
- Utilized Huggingface Transformers, SpaCy, Scikit-Learn, and FlairNLP packages
- Experimented with high-speed model training in Golang

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

Department of East Asian Studies, Johns Hopkins University – Baltimore, MD   *Summer 2020 – Fall 2021*
Computational Research Assistant for Dr. Giovanna Maria Dora Dore
- Built probabilistic models of 4000 news articles with the Causalnex and Scikit-Learn libraries
- Generated plots with Seaborn, Causalnex and Matplotlib to compare news coverage in Hong Kong by year
- Cleaned and combined news datasets with Pandas

Center for Language and Speech Processing (CLSP), Johns Hopkins University – Baltimore, MD   *Summer 2020*
Intern for 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
- Aided in the creation of the largest 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
- Programmed statistical analyses of roadway and car crash data with Pandas and Scikit-learn
- Cleaned and prepared geospatial datasets for machine learning models