

# James “Jamie” Scharf

22 Garey Drive | Chappaqua, NY 10514 | Mobile: (914) 522-2930 | [jscharf8@jhu.edu](mailto:jscharf8@jhu.edu) | <https://github.com/JamesScharf> | [Google Scholar Profile](#)

## SKILLS & QUALIFICATIONS

---

**Clearance:** Secret

**Programming Languages:** Python, C/C++, Java, Go, JavaScript, HTML, Bash/Unix

**Toolkits:** Scikit-Learn, PyTorch, SpaCy, FlairNLP, PySpark, Dask, Matplotlib, OpenNMT, Huggingface Transformers

## EMPLOYMENT & RESEARCH EXPERIENCE

---

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

*Summer 2021*

*Thesis Civic Digital Fellow, Data Science*

- Created a multilingual in-house annotation platform for Named Entity Recognition (NER) and Text Classification with customized neural and non-neural models in the Huggingface Transformers, SpaCy, Scikit-Learn, and FlairNLP packages
- Bootstrapped ML models in 400 languages using word alignment and projection

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

*Spring 2020 & Spring 2021*

*Head Course Assistant for Information Retrieval and Web Agents*

- Assistant and grader 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

**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 and Matplotlib to compare news coverage in Hong Kong by year

**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

**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

## 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, machine translation, and parallel programming
- GPA:** 3.56 (General Honors)

## RECENT PUBLICATIONS & PRESENTATIONS

---

McCarthy, **Scharf**, Dore. “Two Lenses on the Fragrant Harbor: Differences in Western and Hong Kong-based Reporting on the 2019-2020 Protests,” [LaTeCH-CLfL Proceedings](#), November 2021.

McCarthy, **Scharf**, Dore. “A Mixed-Methods Analysis of Western and Hong Kong-based Reporting on the 2019-2020 Protests,” [TADA](#), October 2021.

**Scharf**, McCarthy, Dore. “Evolution and Bias: News Portrayal of Civil Unrest in Hong Kong, 1998-2020,” [PaCSS](#), August 2021.

**Scharf**, McCarthy, Dore. “Characterizing News Portrayal of Civil Unrest in Hong Kong, 1998-2020.” [CASE](#), August 2021.