Yennie Jun

yennie.jun@gmail.com | yenniejun.com | github.com/yenniejun | +44 07788700924

Education

Oxford Internet Institute, University of Oxford — Oxford, United Kingdom
MSc in Social Data Science — *Oct. 2020 - Aug. 2021*

Tufts University — Medford, MA, USA
BS in Computer Science and History, Minor in Music — *Sep. 2013 – May 2017*

Work Experience

AI and Data Science Research Fellow — Remote
UN Global Pulse, New York Lab — *Apr. 2021 – Present*
- Trace discourse shifts in African radio transcriptions using temporal word embeddings and semantic networks
- Operationalize textual analytics tools to aid public health officials monitor the COVID-19 infodemic

Project Co-Lead — Oxford, United Kingdom
Oxford Artificial Intelligence Society — *Oct. 2020 – Present*
- Lead student-driven projects researching biases in NLP systems and multimodal detection of hateful memes

Big Data Researcher — Seoul, South Korea
Big Data Studies Lab, Seoul National University — *Nov. 2019 – Present*
- Gathered, cleaned, modeled, and analyzed COVID-19 news data across 6 countries, 3 languages
- Performed named-entity recognition, resolution, and disambiguation on historical Korean civil-service figures

Software Engineer — Redmond, WA, USA
Microsoft (Education team) — *Sep. 2017 – Sep. 2019*
- Full-stack development for Assignments Service (a Learning Management System within Microsoft Teams)
- Led integration with Immersive Reader, a tool for empowering learners to improve reading and writing skills
- Maintained service health by monitoring telemetry dashboards and investigating high-severity alerts

Teaching Assistant — Medford, MA, USA
Tufts University — *Aug. 2014 – Dec. 2016*
- Led weekly labs on computer science concepts and skills for introductory data structures course
- Planned future projects and assignments with other lab leaders and participated in writing weekly labs

Current Projects & Working Papers

How True is GPT-2: An Empirical Analysis of Intersectional Occupational Biases. To analyze biases in large language models, GPT-2 is prompted to complete sentences related to occupation for intersectional subjects. Predictions are compared to ground-truth US labor data. (*ArXiv preprint*)

Memes in the Wild: Assessing the Generalizability of the Hateful Memes Challenge Dataset. To assess the benchmark datasets commonly used to classify harmful memes on the Internet, Facebook's Hateful Memes Challenge dataset is evaluated using both unimodal and multimodal models. (*ArXiv preprint*)

Data Surveillance and Biocitizenship in the COVID-19 Pandemic: Digital Contact-tracing in South Korea, Hong Kong, Singapore, and Taiwan. The privacy implications of digital contact tracing in East Asian countries during the COVID-19 pandemic is assessed in news media through topic modeling and network analysis.

COVID-19 Texting Service. Texting service providing COVID-19 statistics and common queries for those without broadband access. Working project with Silicon Harlem, NYC.

Technical Skills

Languages: Python, JavaScript, React.js, Node.js/Express.js, SQL, C/C++/C#, HTML/CSS
Libraries: matplotlib, networkx, nltk, numpy, pandas, pytorch, seaborn, scikit-learn
Technologies: conda, Dash/Plotly, Flask, Git, npm, Postman, Twilio, VS Code

Activities & Leadership

Oxford Artificial Intelligence Labs, Researcher (2020-present)
University of Washington Taekwondo Team, Assistant Instructor (2017-2019)
Puget Sound Symphony Orchestra in Seattle, Flutist (2017-2019)