Rui Ouyang, PhD

Machine Learning Researcher & Data Scientist

For my PhD research, I applied **machine learning and natural language processing** models to fight human trafficking. I love the creative side of engineering and research. I have **8+ years** of technical expertise gained from applications ranging from genomics to robotics and more.

Contact

nrobot.dev

nouyang@alum.mit.edu

nrobot

Tech Stack

- Altair
- Bash
- BeautifulSoup
- Flask
- **Git**
- Linux
- **Jupyter**
- **Matplotlib**
- NLTK
- **Python**
- PyTorch
- **Pandas**
- **scikit-learn**
- Scipy
- SpaCy
- SQLite

Education

Harvard University
PhD in Computer Science — 2017 ~ 2023 · 6 yrs

Harvard University
MS in Computer Science — 2017 ~ 2020 · 2 yrs

Massachusetts Institute of Technology
BS in Mechanical Engineering — 2009 ~ 2013 · 4 yrs

Research Experience

MIT Sloan Applied Economics Group — Cambridge, MA
Trafficking in the Illicit Massage Industry — 2020 ~ now · 2yrs

I trained NLP models to research the U.S. illicit massage industry (IMI). Collaborators include IBM Technology for Good, The Network Team, the Wisconsin Department of Justice, and Traffik Analysis Hub.

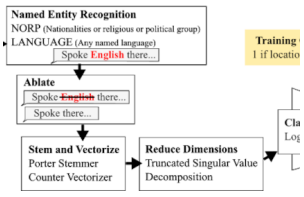
MIT Perceptual Science Group — Cambridge, MA
Touch Sensing for Robot Hands — 2020 · 1yr

Developed tactile sensor for buried objects. My work in electromechanical prototyping and image-based object detection and shape recognition with **PyTorch** led to an MIT news article & second-author publication in ISER 2020.

Harvard Biorobotics Lab — Cambridge, MA
Vision-Based Force Sensing — 2018 · 1yr

Designed and characterized novel vision-based force sensor using arucoTags and **OpenCV**, leading to first-author publication in ICRA 2020. The sensor software and hardware design files are open-sourced.

Publications



Machine Learning for Tangible Effects: Natural Language Processing for Uncovering the Illicit Massage Industry

Harvard Uni. Department of Computer Science, Doctoral Thesis, 2023
Rui Ouyang

- thesis on arxiv - defense video - defense slides - invited talk

The United States illicit massage industry (IMI) is a multi-billion dollar industry. By creating datasets with three publicly-accessible websites and appyling NLP techniques such as named entity recognition, bag-of-words and Word2Vec, I derive insights into the labor pressures and language barriers that employees in the IMI face, as well as the

Robotics Stack

- Arduino/ microcontrllers
- Diptrace (PCB Design)
- OpenCV
- ROS
- Solidworks
- 3D Printer
- Lasercutter
- Mill/Lathe

Creative Stack

- GIMP (Photo Editing)
- Inkscape (Vector Art)
- Jekyll
- Lightworks (Video Editing)
- Wordpress

Hobbies

- Sailing on the ocean

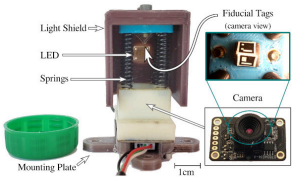
income, demographics of, and social pressures affecting sex buyers. Committee: Profs. Roberto Rigobon, Finale Doshi-Velez, David Parkes.

Digger Finger: GelSight Tactile Sensor for Object Identification Inside Granular Media

International Symposium on Experimental Robotics (ISER), 2020
Radhen Patel, **Rui Ouyang**, Branden Romero, Edward Adelson

- arxiv - slides - site - MIT News

I worked with Radhen Patel to prototype a robotic sensor specialized for finding objects buried in sand. We created a wedge-shaped sensor with a vibrator motor and miniaturized the Gelsight technology so as to reduce force in traversing sand. Featured on MIT News.



Low-Cost Fiducial-Based 6-Axis Force-Torque Sensor

International Conference on Learning Representations (ICLR), 2020
Rui Ouyang, Robert Howe

- arxiv - slides - site

A novel miniature six-axis force/torque sensor. The sensor is low-cost (<$50) compared to other sensors which can cost up to tens of thousands of dollars (e.g. some ATI sensors). The sensing component itself uses simple webcam and a printed paper tag, allowing for easy customization and use. Design files and code are open-source.

Industry Experience

Scotiabank
Doctoral Intern

Toronto, Canada (Remote)
2021 ~ 2022 · 2yrs

Designed **agent-based models** to generate synthetic data to benchmark anomaly detection algorithms such as **Gaussian Mixture Models** used to flag transactional data and accounts for counter-trafficking work.

Curoverse
Researcher

Somerville, MA
2014 ~ 2015 · 1yr

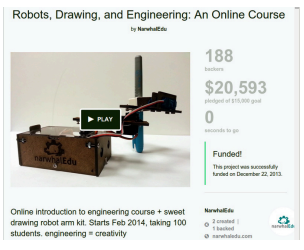
Developed privacy-respecting **Python Flask** webapp to share genomic data in the Global Alliance for Genomics and Health (GA4GH). Designed Arvados platform documentation with Boostrap CSS as part of Agile team.

NarwhalEdu
Founder, CEO

Cambridge, MA
2013 ~ 2014 · 1yr

Founded startup to expand access to robotics classes by combining online classes with low-cost hardware kits. Accepted into MIT Global Founders' Skills Acclerator. My **entrepreneurial and grant-writing skills** led to >$45k in funding and sales in our first year.

Startup



NarwhalEdu: Creativity and Engineering

· 2014

- more details

Founded an educational company, NarwhalEdu, to highlight the creative side of engineering. Created "Creativity in Engineering" class on EdX with a custom robot arm kit, achieving $20k of pre-orders. Product (contract-manufactured lasercut chassis kit) delivered on-time and on-budget to students around the world.