

Ken Kato

kenkato.info • ken.kato@mail.utoronto.ca • github.com/kkatodus • [linkedin.com/in/kkatodus](https://www.linkedin.com/in/kkatodus) • JPN Citizen/GER PR/CAN Work Permit

Education

University of Toronto

09.2019 - 2020, 2021 - 2025

Bachelor of Applied Science in Engineering Science, Major in Machine Intelligence

Major GPA:3.64

Work Experience

AI/Robotics Researcher | LSY Lab - Technical University of Munich

Munich, 05.2024 - 08.2024

- Worked towards integrating **LLMs** into robotic controls to enable semantically nuanced behaviour of human assisting robots
- Developed a **simulation** environment in **dm_control** for the Stretch 3 platform to test out the deployment of **RT1** model on the platform
- Integrated **RT1** model into a **ROS2 node** to control the physical **Stretch 3** to conduct picking up tasks with **70% success rate**
- Designed soft-material robotic grippers and iterated on the design to successfully create a motor actuated soft gripper

SWE Intern | Hellofresh Canada

Toronto, 05.2023 - 05.2024

- Accelerated the development of the rapidly expanding **web platform** implemented in **React**, **NextJS** and **Typescript** for HelloFresh and its sub-brands including ChefsPlate, Factor, GreenChef, and EveryPlate, enhancing global user access
- Engineered and executed end-to-end tests with **Cypress** and developed unit tests with **React Testing Library** to ensure software performance.
- Established monitoring protocols using **Honeycomb**, implemented **Google Analytics tracking**, and led **A/B testing** initiatives with **Statsig** and **Optimizely** to optimise user experience and operational efficiency

NLP Researcher | U of T research with Dr. Christopher Cochrane and Prof. Raed Saqur

Toronto, 10.2023 - 04.2024

- Explored the use of **LLM** in the analysis of parliamentary speeches of Japanese legislators, predicting their positions on the political spectrum
- Analysed speeches of **464 parliamentary representatives** totalling **129,859** speech segments to infer their political stances
- Fine-tuned a **SBERT classifier** with self-labelled dataset composed of **1,439 speech segments** and used **BERTopic** to gain insight to clusters
- Visualised the resulting data by reducing the dimensionality of embeddings by **UMAP** and plotting each representative's ideological position
- Published a paper as **first-author** to **PolMeth2024** conference and presented a poster presentation among top researchers in the field

NLP Researcher | U of T/Vector Institute research with Prof. Raed Saqur

Toronto, 10.2023 - 04.2024

- Worked as a NLP researcher under the supervision of Prof. Raed Saqur to apply LLMs to stock price prediction tasks
- Curated a filtered, cleaned dataset out of **10 years** of financial headlines composed of headlines from NYT, WSJ and Reuters and financial market data totalling **2,111 data points** which can be used for **supervised finetuning of LLMs** or alignment methods such as **RLHF**

Research Fellow | Citizen Lab

Toronto, 05.2022 - 05.2023

- Conducted a comprehensive analysis of **ensorship** mechanisms on Chinese search engines to find **60,000 unique censorship** rules across 8 major **Chinese ISPs** by developing a **Python** algorithm using **Selenium** to scrape the search engines of the websites
- Discovered novel censorship rules unknown previously to the public which we named **hard censorship** and **soft censorship**
- Developed and optimised an iterative **search algorithm** to refine and identify censored keyword combinations using text data from independent news sources and detecting key word combinations in the text that is being censored
- Gathered and synthesised data on censored keywords, revealing a consistent pattern of censorship across various platforms
- Published a **50 page** report on the research which was featured on the **New York Times** which gained global attention

Software Engineer Intern | Kozo Keikaku Engineering Inc.

Tokyo, 08.2020 - 06.2021

- Contributed to the web application of the company coding both the backend written in **flask** and frontend written in **HTML**, **CSS** and **JS**
- Data analysis using tweets and evaluating the trend visible on the platform Used natural language analysis tools such as **Mecab** and **Gensim** and data analysis modules such as **pandas**, **matplotlib** and **scipy**
- Led **fluid simulation** tasks in **SolidWorks Fluid Simulation** in order to evaluate the performance of inkjet printers of industry leading clients
- Implemented **deep reinforcement learning models** such as **A2C**, **REINFORCE** and **DQN** in python to act in a financial market simulation

Personal Projects

KOKKAI DOC | Project for Political Transparency in Japan through NLP, Data Viz and Web Dev

05.2022 - Present

[Project Link](#) [Website link](#) [Demo video](#)

- Established an innovative web service dedicated to aggregating and disseminating Japanese parliamentary speech and voting data
- Platform currently attracts **≈1000** unique users per month and youtube videos relating to the platform have been watched **≈60,000** times in total
- Employed **LLM** embeddings to perform quantitative analysis of ideologies of politicians by placing them on the political spectrum
- Utilised **Selenium** for **data scraping** and Python for processing to ensure robust and reliable data acquisition for the platform
- Enhanced a Hugging Face Large Language Model by **fine-tuning** it with self-labelled data to accurately classify opinion-based sentences
- Developed the frontend of the project using the **React** framework, integrating the **deck.gl** package for advanced **geospatial data visualisation**

Fullstack diary application | Web application using Django, React, PsqI

06.2021 - 07.2021

[Project Link](#) [Demo video](#)

- Designed and launched a **fullstack** web app enabling users to manage their diaries, including **CRUD** operations and picture uploads
- Engineered a **RESTful** API using **Django** and **psql** and developed the frontend with **React** to enhance user interaction and functionality

Publications

[L\(u\)PIN: LLM-based Political Ideology Nowcasting](#) | [PolMeth 2024 Poster Presentation](#)

Ken Kato, Annabelle Purnomo, Christopher Cochrane, Raed Saqur

[Missing Links - A comparison of search censorship in China](#) | [Featured on New York Times](#)

Jeffrey Knockel, Ken Kato, Emile Dirks [New York Times Article](#)

[NIFTY Financial News Headlines Dataset](#)

Raed Saqur, Ken Kato, Nicholas Vinden, Frank Rudzicz

Awards

[ESROP-GLOBAL Award](#)

- Received funding worth 7000 CAD to conduct research at the Technical University of Munich as an AI researcher over the summer of 2024

[UofT IE+ Award](#)

- Received funding worth 2500 CAD to conduct research at the Technical University of Munich as an AI researcher over the summer of 2024

[ESROP-E4TW Award](#)

- Received funding worth 7000 CAD to conduct research at the Citizen Lab over the summer break of 2022