

MIKE ZHANG

+31 6 52895785 ◊ Copenhagen, Denmark

mikz@itu.dk ◊ [linkedin.com/in/jjzha](https://www.linkedin.com/in/jjzha) ◊ github.com/jjzha

EDUCATION

IT University of Copenhagen / Copenhagen, DK
Ludwig Maximilian University of Munich, Research Visitor Munich, DE
PhD, Natural Language Processing Dec. 2023

Thesis: Natural Language Processing for Labor Market Information. Advised by prof. [Barbara Plank](#).
Award: **2,000 USD** for an **Outstanding Paper Award** (ML Evaluation Standards Workshop, ICLR 2022).

University of Groningen Groningen, NL
MA, Information Science Jan. 2020
Thesis: The Effect of Translationese in Machine Translation Test Sets.

The Chinese University of Hong Kong Hong Kong, HK
Study abroad, Computer Science May 2018
Award: **1,000 EUR** for studying abroad outside of Europe, one awardee per semester.
Courseswork: Computer Networks, Java Programming, Politics.

University of Groningen Groningen, NL
BSc, Information Science Jun. 2018
Thesis: Labeled Bilexical Dependencies for Deception Detection.

WORK EXPERIENCE

National University of Singapore Singapore, SG
Research Intern Jan. 2023 – Jul. 2023

- Research visit at the NUS, WING Group. Advised by prof. [Min-Yen Kan](#). Joint research with respect to NLP and IR related to job descriptions and related text sources.

Dataprovider.com Groningen, NL
Data Engineer Jan. 2020 – Aug. 2020

- Created requested features into the proprietary search engine by maintaining code quality with PEP8 standards, basic unit- and integration tests, Git CI/CD, and code reviews;
- Spearheaded improvement of the preprocessing steps for structuring 72M unstructured company addresses with Python and Google's Geocoding API;
- Implemented a classifier predicting company SIC codes via character-level features using Scikit-Learn, improved predictions by 70% accuracy.

Research Intern Aug. 2019 – Dec. 2019

- Developed five text generation algorithms in Python for the generation of undiscovered country code top-level domains for five countries, found ~15K undiscovered domain names/hr;

PUBLICATIONS

Latest work can be found at [Google Scholar](#).

SERVICES

Reviewer NAACL SRW (2022), ARR (2021, 2022), W-NUT (2021), CoNLL (2021), ACL (2019)
Volunteer EMNLP (2021), CLIN (2019)

SKILLS

Programming Python, Java, R, HTML, CSS
Tools Git, Doccano, Scikit-Learn, PyTorch
Hard Skills Scientific Writing, Research Management, Public Speaking, Data Analysis
Soft Skills Time Management, Teamwork, Communication, Problem Solving, Leadership, Accountability
Languages Dutch (Native), English (Fluent), Spanish (Basic), Mandarin (Basic), Danish (Basic)