YU CAO

Computational Linguist

⚲ Berlin, Germany   @ gencaetia@gmail.com   ⚲ pterosdiacos.github.io   in www.linkedin.com/in/yu-cao-pdiacos

PROFESSIONAL PROFILE

Linguistics   Specialized in computational linguistics, semantics, syntax, discourse analysis. Strong critical-thinking and problem-solving skills.

Computation   Expertise in NLP/NLU modeling and machine learning algorithms. 15+ years of experiences in algorithms and data structures.

EXPERIENCES

Applied Researcher
eBay

📅 11/2021 – present   📍 Berlin, Germany; San Jose, CA

- Work at Content, Item and Inventory Understanding Team.
- Build machine learning solutions that improve content understanding and drive search quality.

Ontology Linguist
Facebook via PRO Unlimited

📅 08/2019 – 10/2019   📍 Redmond, WA

- Worked at Assistant Cross-Functional Group.
- Prototyped cross-domain semantic parsing models for question answering.
- Automated inferences in *OWL 2* ontology knowledge graphs.

Teaching Assistant
Rutgers University

📅 09/2018 – 08/2020   📍 New Brunswick, NJ

- Courses taught: *Intro to Linguistic Theory, Invented Languages*.

Research Assistant
The Chinese University of Hong Kong

📅 12/2014 – 07/2015   📍 Hong Kong, China

- Worked at Center for Sign Linguistics and Deaf Studies.
- Provided linguistic expertise in developing Chinese grammatical knowledge assessment software.

EDUCATION

Rutgers University
Ph.D. in Linguistics

📅 09/2016 – 10/2021   📍 New Brunswick, NJ

The Chinese University of Hong Kong
M.A. in Chinese Linguistics & Language Acquisition

📅 09/2013 – 06/2014   📍 Hong Kong, China

Sichuan Normal University
B.A. in Teaching Chinese as a Second Language

📅 09/2009 – 06/2013   📍 Chengdu, China

PUBLICATIONS

- Semantic Similarity Based Label Augmentation for Visual Classification
  Tiny Papers Track at ICLR 2023.
  👤 **Yu Cao**.

- Plurality and quantification in graph representation of meaning (⚲ code)
  Doctoral dissertation on arXiv:2112.06448.
  👤 **Yu Cao**.
  Advisor: Simon Charlow.

- Automated fact-value distinction in court opinions
  European Journal of Law and Economics 2020.
  👤 **Yu Cao**, Daniel Chen & Elliot Ash.
  Automated classification of fact vs. value statements in written judicial decisions.

- Investigating BERT's knowledge of language: Five analysis methods with NPIs
  Proceedings of EMNLP 2019.
  👤 Alex W., **Yu Cao** & 14 others.
  Supervisor: Sam Bowman.
  Role: constructed and conducted the main transfer learning experiments.

SKILLS

Technical

- Python, C, Haskell, SQL, PyTorch, Tensorflow
- Linguistic analysis, symbolic logic, statistics

ML/NLP Coursework

- Natural language processing ⚲
  📅 09/2020   🏛 HSE@Coursera
- Bayesian methods for machine learning ⚲
  📅 08/2020   🏛 HSE@Coursera
- Introduction to deep learning ⚲
  📅 07/2020   🏛 HSE@Coursera
- Machine learning
  📅 07/2020   🏛 Stanford@Coursera
- Seminar: Linguistic knowledge in reusable sentence encoders
  📅 05/2019   🏛 NYU
- Graph formalisms for meaning representations
  📅 07/2018   🏛 NASSLLI@CMU
- NLU & computational semantics
  📅 05/2018   🏛 NYU

Languages

ENG (fluent)   CHN (native)   JPN (JLPT N1)