

Ruixiang Cui

Lyngbyvej 2, Copenhagen 2100, Denmark

E-mail: rc@di.ku.dk

EDUCATION

University of Copenhagen

PhD in Computer Science, Natural Language Processing

Supervisors: Daniel Hershcovich, Anders Søgaard

Jan. 2021 - June 2024 (expected)

University of Copenhagen

MSc in IT and Cognition

2018-2020

Jilin University

BA in Hispanic Philology

2014-2018

ACADEMIC STAYS

Stanford University

Guest Researcher

Host: Christopher Potts, Stanford NLP Group

Sept. 2023 - Nov. 2023

WORK EXPERIENCE

Microsoft Research

Research Intern

Jan. 2023 - June 2023

- Led to establish a bilingual benchmark for testing the general abilities of large language models (mainly GPT-4) in human exams. The benchmark has been acknowledged for model evaluation by companies such as Microsoft, Bytedance and SenseTime.
- Worked on building an automatic long-content generation system with GPT-4.

Industrial and Commercial Bank of China

Data Analyst Intern

June 2016 - Sept. 2016

- Assisted in building an information extraction system for letters of credit transactions. Created APIs to extract numerical data from financial and commercial documents in Python.
- Used Excel to organize and analyze commercial invoices and bills of lading.

RESEARCH INTERESTS

benchmarking and evaluation of large language models, multilingual and multi-cultural analysis, compositional generalization, knowledge-base semantic parsing, and natural language inference.

PUBLICATIONS

- [1] Wanjun Zhong*, **Ruixiang Cui***, Yiduo Guo, Yaobo Liang, Shuai Lu, Yanlin Wang, Amin Saied, Weizhu Chen, Nan Duan. 2023. *AGIEval: A Human-Centric Benchmark for Evaluating Foundation Models*. arXiv.
- [2] **Ruixiang Cui**, Seolhwa Lee, Daniel Hershcovich, Anders Søgaard. 2023. *What does the Failure to Reason with “Respectively” in Zero/Few-Shot Settings Tell Us about Language Models?*. Annual Meeting of the Association for Computational Linguistics (ACL).
- [3] Nikolaus Schrack, **Ruixiang Cui**, Hugo A. López, Daniel Hershcovich. 2022. *Can AMR Assist Legal and Logical Reasoning?*. Findings of Conference on Empirical Methods in Natural Language Processing (EMNLP).

- [4] **Ruixiang Cui**, Rahul Aralikkatte, Heather Lent, Daniel Hershcovich. 2022. *Compositional Generalization in Multilingual Semantic Parsing over Wikidata*. Transactions of the Association for Computational Linguistics (TACL, oral presentation at EMNLP).
- [5] **Ruixiang Cui**, Daniel Hershcovich, Anders Søgaard. 2022. *Generalized Quantifiers as a Source of Error in Multilingual NLU Benchmarks*. North American Chapter of the Association for Computational Linguistics (NAACL, oral).
- [6] Stephanie Brandl, **Ruixiang Cui**, Anders Søgaard. 2022. *How Conservative are Language Models? Adapting to the Introduction of Gender-Neutral Pronouns*. North American Chapter of the Association for Computational Linguistics (NAACL).
- [7] Daniel Hershcovich, Stella Frank, Heather Lent, Miryam de Lhoneux, Mostafa Abdou, Stephanie Brandl, Emanuele Bugliarello, Laura Cabello Piqueras, Ilias Chalkidis, **Ruixiang Cui**, Constanza Fierro, Katerina Margatina, Phillip Rust, Anders Søgaard. 2022. *Challenges and Strategies in Cross-Cultural NLP*. Annual Meeting of the Association for Computational Linguistics (ACL).
- [8] **Ruixiang Cui**, Daniel Hershcovich. 2021. *Great Service! Fine-grained Parsing of Implicit Arguments*. International Conference on Parsing Technologies (IWPT).
- [9] **Ruixiang Cui**, Daniel Hershcovich. 2021. *Meaning Representation of Numeric Fused-Heads in UCCA*. ACL Workshop on Understanding Implicit and Underspecified Language.
- [10] **Ruixiang Cui**, Daniel Hershcovich. 2020. *Refining Implicit Argument Annotation for UCCA*. COLING Workshop on Designing Meaning Representations.
- [11] Ofir Arviv, **Ruixiang Cui**, Daniel Hershcovich. 2020. HUJI-KU at MRP 2020: *Two Transition-based Neural Parsers*. CoNLL Shared Task on Cross-Framework Meaning Representation Parsing.

SKILLS

Programming Languages	Python, Java, SQL, SPARQL, LaTeX, Visual Basic, Perl
Frameworks and Tools	PyTorch, TensorFlow, Keras, SciPy, Scikit-learn, NumPy, Pandas, Matplotlib, NLTK, OpenCV, Git, grep, Excel, Tableau
Natural Languages	Mandarin Chinese (Native), English (Full professional), Spanish (Professional), Danish (Limited working), Portuguese (Limited working), Italian (Elementary)

HONORS & AWARDS

Excellent Student Leader & Second Prize Scholarship	2016
First Prize Scholarship	2015

TEACHING

Natural Language Processing, University of Copenhagen	2022
Data Science (Machine Learning), University of Copenhagen	2021
Data Science (Database Design), University of Copenhagen	2021
Spanish Grammar, Jilin University	2018

SUMMER & WINTER SCHOOLS

University of Oxford	Aug. 2022
Machine Learning Summer School	

HOBBIES

Spanish Guitar, Russian Literature, Brazilian Music, Electronic Music Producer, World Heritage Traveller