

# Ruixiang Cui

E-mail: rc@di.ku.dk ◇ Homepage: ruixiangcui.github.io

Google Scholar: 3M-1U2wAAAAJ ◇ GitHub: ruixiangcui ◇ LinkedIn: ruixiang-cui

## EXPERTISE

---

Large language models (instruction tuning, prompt engineering, data curation and evaluation), deep learning-based NLP tools, cloud computing stacks (Azure, GCP, AWS), SQL databases, knowledge-base semantic parsing

## EDUCATION

---

<b>University of Copenhagen</b> PhD in Computer Science, Natural Language Processing Advisors: Daniel Hershcovich, Anders Søgaard	2021-2024( <i>expected to graduate in July</i> )
<b>Stanford University</b> Visiting Researcher, Stanford NLP Group	2023
<b>University of Oxford</b> Machine Learning Summer School	2022
<b>University of Copenhagen</b> MSc in IT and Cognition	2018-2020
<b>Jilin University; Complutense University of Madrid</b> BA in Hispanic Philology	2014-2018

## EXPERIENCE

---

<b>Microsoft</b> <i>Research Intern</i>	01/2023-06/2023
<ul style="list-style-type: none"><li>• Led <i>AGTEval</i>, a benchmark for assessing the general capabilities of LLMs. Orchestrated the design and execution of data management pipelines, encompassing data collection, refinement, and analysis.</li><li>• Devised and led the creation of a human workflow platform utilizing LLM for generating long-form content in AI + HCI situations, offering a customizable and scalable solution.</li></ul>	
<b>University of Copenhagen</b> <i>Teaching Assistant</i>	09/2021-12/2022
<ul style="list-style-type: none"><li>• Taught and supervised students in three courses: Natural Language Processing (MSc), Machine Learning (BSc) and Database Design (BSc).</li></ul>	
<b>Industrial and Commercial Bank of China</b> <i>Data Analyst Intern</i>	06/2016-09/2016
<ul style="list-style-type: none"><li>• Crafted and deployed an information extraction system for letters of credit transactions. This system significantly streamlined data processing, enhancing accuracy and efficiency in financial operations.</li></ul>	

## SKILLS

---

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

## HONORS & AWARDS

---

Excellent Internship Award  
Excellent Student Leader & Second Prize Scholarship  
First Prize Scholarship

Microsoft Research 2023  
Jilin Univeristy 2016  
Jilin Univeristy 2015

## SELECTED PUBLICATIONS (CITATIONS: 300; H-INDEX: 7)

---

- [1] Wanjun Zhong\*, **Ruixiang Cui\***, Yiduo Guo, Yaobo Liang, Shuai Lu, Yanlin Wang, Amin Saied, Weizhu Chen, Nan Duan. 2024. *AGIEval: A Human-Centric Benchmark for Evaluating Foundation Models*. Findings of North American Chapter of the Association for Computational Linguistics (NAACL).
- [2] Yong Cao, Yova Kementchedjhieva, **Ruixiang Cui**, Antonia Karamolegkou, Li Zhou, Megan Dare, Lucia Donatelli, Daniel Hershcovich. 2024. *Cultural Adaptation of Recipes*. Transactions of the Association for Computational Linguistics (TACL).
- [3] **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).
- [4] 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).
- [5] **Ruixiang Cui**, Rahul Aralikatte, 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).
- [6] **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).
- [7] 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).
- [8] 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).

## HOBBIES

---

Literature, Popular Music Criticism, Travelling, Philosophy, Art, Biking, Running