# Bohui Zhang


**Contact Information**

bohui.zhang@kcl.ac.uk
https://bohuizhang.github.io/


**Research Interests**

My research interests mainly lie in knowledge graphs, natural language processing, and explainable AI—particularly mechanistic interpretability. I focus on automatic and explainable knowledge graph construction, natural language processing for knowledge graph development, and the use of multi-agent systems in this context.

**Academic Background**

**King's College London (KCL)** 2022 - 2026
Ph.D. in Computer Science
Supervisors: Prof. Elena Simperl, Dr. Albert Meroño Peñuela

**University of Southern California (USC)** 2020 - 2021
M.S. in Applied Data Science

**University of Waterloo (UWaterloo)** 2015 - 2019
B.S. in Materials and Nanosciences, *Dean's Honours List*

**Beijing Jiaotong University (BJTU)** 2015 - 2019
B.Eng. in Nanomaterials and Nanotechnology

**Work Experience**

**Information Sciences Institute**, Student Researcher 09/2021 - 12/2021

- Supervisor: Dr. Filip Ilievski
- Investigated the feasibility of enriching Wikidata with structured data sources from the linked open data (LOD) cloud.
- Proposed a method that consists of several steps: gap detection, external graph selection, schema alignment, knowledge retrieval, and validation, implemented the procedure using the Knowledge Graph Toolkit (KGTK).
- Evaluated the method on enriching Wikidata with two LOD sources: DBpedia and Getty Vocabularies. The experiments showed that the LOD-based method can enrich Wikidata with millions of new high-quality statements in a short time.

**Alibaba Cloud**, Machine Learning Intern 05/2021 - 08/2021

- Mentor: Jingjun (Alvin) Chu
- Worked on a Neural Architecture Search (NAS) system for optimizing models in search space defined by ProxylessNAS, used on image classification and feature extraction tasks based on dataset collecting from group's retail sector.
- Improved the model training process using knowledge distillation and improved the optimal model architecture searching process in various hardware environments using policy gradient algorithm based on target accuracy, FLOPs and latency.
- The optimal models deployed on terminal machines achieved model compression for more than 60% decrease on FLOPs while improving rank1 and rank6 compared with state-of-the-art MobileNetV2 models and keeping the top1 accuracy above 98%.

**Publications** [8]

8. Bohui Zhang, Yuan He, Lydia Pintscher, Albert Meroño-Peñuela, Elena Simperl, Schema Generation for Large Knowledge Graphs Using Large Language Models. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, 2025
7. Bohui Zhang, Elisavet Koutsiana, Yihang Zhao, Albert Meroño-Peñuela, and Elena Simperl, Trustworthy Knowledge Graphs: Practices and Approaches. In *Handbook on Neurosymbolic AI and Knowledge Graphs*. IOS Press, 2025
6. Elisavet Koutsiana, Johanna Walker, Michelle Nwachukwu, Bohui Zhang, Albert Meroño-Peñuela, Elena Simperl, Knowledge Prompting: How Knowledge Engineers Use Generative AI, In *Journal of Web Semantics*, 2025
5. Bohui Zhang, Albert Meroño-Peñuela, and Elena Simperl, Towards Explainable Automated Knowledge Engineering with Human-in-the-loop, In *Semantic Web Journal*, 2025
4. Bohui Zhang, Valentina Anita Carriero, Katrin Schreiberhuber, Stefani Tsaneva, Lucía Sánchez González, Jongmo Kim, Jacopo de Berardinis, OntoChat: a Framework for Conversational Ontology Engineering using Language Models, In *Extended Semantic Web Conference*, 2024
3. Bohui Zhang, Ioannis Reklos, Nitisha Jain, Albert Meroño Peñuela, Elena Simperl, Using Large Language Models for Knowledge Engineering (LLMKE): A Case Study on Wikidata, In *Knowledge Base Construction from Pre-trained Language Models Workshop at International Semantic Web Conference*, 2023
2. Bohui Zhang, Albert Meroño Peñuela, Elena Simperl, Towards Explainable Automatic Knowledge Graph Construction with Human-in-the-loop, In *International Conference on Hybrid Human-Artificial Intelligence* (HHAI), 2023
1. Bohui Zhang, Filip Ilievski, Pedro Szekely, Enriching Wikidata with Linked Open Data, In *Wikidata Workshop co-located with International Semantic Web Conference*, 2022

**Teaching Assistant**

5CCSAKNE Knowledge Engineering KCL

- Module leader: Dr. Albert Meroño Peñuela, Prof. Elena Simperl
- Developed course materials, including lectures and lab sheets, on knowledge graphs, knowledge engineering, and large language models; led lab sessions.
- Semesters: 2024-25 Semester 2

7CUSMNDA Network Data Analysis KCL

- Module leader: Dr. Albert Meroño Peñuela
- Designed and delivered coding lab sessions, topics covered included graph theory, spatial and social network analysis, graph embedding, and semantic web.
- Semesters: 2022-23 Semester 2, 2023-24 Semester 2, 2024-25 Semester 2

5CCS2FC2 Foundations of Computing II KCL

- Module leader: Dr. Christopher Hampson
- Delivered lab sessions on algorithm problems, topics covered included P/NP, SAT solving, approximation, linear programming, and probabilistic algorithms.
- Semesters: 2022-23 Semester 1

**Awards**

| | |
|---|---|
| NMES Enterprise & Engagement Partnerships Fund | KCL, 2023 |
| Graduation Dean's Honours List | UWaterloo, 2019 |
| Waterloo-Beijing Jiaotong University Tuition Award | BJTU, 2016, 2017, 2018 |
| Excellence Scholarship of Academic Activities | BJTU, 2017 - 2018 |
| Excellence Scholarship of Social Activities | BJTU, 2015 - 2016 |

**Academic Service**

**Organizer** for ELMKE Workshop series, LM-KBC Challenge series, Knowledge Prompting Hackathon 2023.
**Reviewer** for Information Processing & Management, Semantic Web, SEA@NIPS 2025, ISWC 2024, ACM CHI 2023.
**Member** of Knowledge Graphs Interest Group at the Alan Turing Institute.

**Skills**

**Programming Languages:** Python, Java, JavaScript, MATLAB
**Frameworks & Platforms:** Docker, LangChain, PyTorch, transformers, KGTK, OpenAI Agents
**Semantic Web Tech Stacks:** OWL, RDF, SHACL, ShEx, SPARQL, PROV
**Databases:** MongoDB, MySQL, Neo4j