

Roya Yunjie He

+49(0)152-0531-3060 | heyunjieroya@gmail.com
<https://royahe.github.io/>



EDUCATION

Max Planck Institute for Intelligent Systems & University of Stuttgart

Artificial Intelligence Doctor Full-time

Jun 2022 - Present

Stuttgart, Germany

- Supervisor: Prof. Dr. Steffen Staab, Prof. Dr. Evgeny Kharlamov
- Research direction: Question Answering over Knowledge Graphs, Knowledge Representation and Reasoning, KG RAG for LLMs, information retrieval, representation learning of structure and semi-structured data.
- Ph. D. Thesis: Querying Incomplete Knowledge Graphs with Embedding
- Achievements: Four first-authored papers accepted by EMNLP, WWW, ECAI, and ISWC; Two patents approved.

University College London

MSc Computational Statistics and Machine Learning, Computer Science Department

Sep 2020 - Sep 2021

London, United Kingdom

- Degree: Distinction Certified with Outstanding Academic Performance
- Courses: Natural Language Processing, Deep Learning, Supervised Learning, Reinforcement Learning, Probabilistic Unsupervised Learning, Approximate Inference.
- In-class Research: Discovering the Effectiveness of Pre-trained Masked Language Models for English Spell Correction

University College London

BSc Economics and Statistics, Statistical Science Department

Sep 2017 - Jun 2020

London, United Kingdom

- Degree: Bachelor (First Class Honours)
- Honors/Awards: International Student Scholarship and Summer Research Scholarship
- Related courses: linear algebra, calculus, matrix, econometrics, macro-microeconomics, statistical probability and inference, Bayesian methods, random systems

PROFESSIONAL EXPERIENCE

SAP SE

Jun 2026 - Dec 2026

PhD Researcher Intern on Knowledge Graphs for Agents, Business Knowledge Graph

Walldorf, Germany

- Conducted large-scale structural and semantic analysis of SAP S/4HANA enterprise knowledge graphs, identifying key characteristics and challenges such as hierarchical schemas, uneven data distribution, context-dependent semantics, and ambiguity across business domains.
- Formulated requirement-fulfilment retrieval to address the limitations of conventional relevance-based retrieval, which may return individually relevant evidence without ensuring that a single answer entity satisfies all user requirements.
- Developed a training-free KG-RAG framework that improves answer reliability by combining LLM-based question decomposition, schema-guided evidence retrieval, candidate-level contextual verification, and explicit requirement fulfilment, enabling more precise and explainable question answering on semi-structured enterprise knowledge graphs

Bosch Center for Artificial Intelligence

Mar 2023 - Apr 2026

Doctoral Researcher, Neural-Symbolic AI & NLP

Stuttgart, Germany

- Led the research and development of large-scale knowledge graph reasoning algorithms in Bosch industrial scenarios, combining large language models, logical reasoning and embedding methods to build enterprise knowledge-graph question-answering systems.
- In-depth study of the interpretability of the current knowledge graph ML reasoning model, the support for complex query and hybrid search under multimodal data that combines text and graph. The research results successfully promoted 2 patent applications.

HUAWEI Noah's Ark Lab

Jun 2021 - Oct 2021

Research Intern, NLP Group

London, UK

- Collaborated with Dr. Philip Gorinski and Prof. Dr. Pontus Stenetorp to redesign the Hierarchical Graph Networks architecture, integrating a multi-head attention mechanism and enhanced graph convolution operations to strengthen multi-hop reasoning capabilities.
- Engineered an improved Hierarchical Graph Attention model and conducted systematic evaluations on HotpotQA and other large-scale NLP benchmarks, measurable gains in complex multi-step question answering tasks.

Huatai Technology

Jun 2020 - Aug 2020

- Orchestrated end-to-end data workflows leveraging Hive and Sqoop to ingest, transform, and analyze high-volume video traffic datasets from the Eye-Opening App, processing over **1 million records daily** to derive actionable insights on content category popularity and user engagement patterns.
- Translated analytical findings into strategic business recommendations, identifying underperforming video categories and growth opportunities that directly informed content acquisition decisions and contributed to a **15% uplift** in targeted category viewership within one quarter.

Alan Turing Institute

May 2019 - Aug 2019

Research Assistant, Machine Learning

London, UK

- Participate in the development of ML automation toolkit "mlrproba", aiming to build a ML automation toolkit with object-oriented programming language as the main framework with R. Based on the previous mlr3, it is improved and expanded in areas such as Bayesian supervised learning.
- Leading the construction of various basic distributions and the preparation of decorator, as well as connecting existing sampling software such as BUGS and JAGS to the preparation of Monte Carlo methods.

SELECTED PUBLICATIONS

1. [EMNLP Main 2026] **Y. He**, B. Xiong, D. Hernández, Y. Zhu, Y. Wang, E. Kharlamov, and S. Staab, "Logic meets language: Hybrid query answering over incomplete text-labeled graphs"
2. [TMLR, 2026] Y. Brunink, D. Daza, **Y. He**, and M. Cochez, "Counting still counts: Understanding neural complex query answering through query relaxation"
3. [ACL Industry Track, 2026] P. Chaturvedi, Y. Zhu, H. Zhou, D. Zhou, **Yunjie He**, S. Staab, F. Du, J. Tang, and E. Kharlamov, "SCAIR: Schema-conditioned agentic iterative reasoning for enterprise knowledge graphs"
4. [WWW, 2025] **Y. He**, B. Xiong, D. Hernández, Y. Zhu, E. Kharlamov, and S. Staab, "DAGE: DAG query answering via relational combinator with logical constraints"
5. [Handbook on Neurosymbolic AI and Knowledge Graphs, 2025] B. Xiong, M. Nayyeri, M. Jin, **Y. He**, M. Cochez, S. Pan, and S. Staab, "Geometric relational embeddings: Progress and prospects"
6. [NAACL Main, 2025] Y. Zhu, N. Potyka, J. Pan, B. Xiong, **Y. He**, E. Kharlamov, and S. Staab, "Conformalized answer set prediction for knowledge graph embedding"
7. [ECAI Oral, 2024] **Y. He**, D. Hernández, M. Nayyeri, B. Xiong, Y. Zhu, E. Kharlamov, and S. Staab, "Generating SROI ontologies via knowledge graph query embedding learning"
8. [AAMAS, 2024] N. Potyka, Y. Zhu, **Y. He**, E. Kharlamov, and S. Staab, "Robust knowledge extraction from large language models using social choice theory"
9. [ISWC Poster, 2023] **Y. He**, M. Nayyeri, B. Xiong, Y. Zhu, E. Kharlamov, and S. Staab, "Can pattern learning enhance complex logical query answering?"
10. [EMNLP Findings, 2023] Y. Zhu, N. Potyka, M. Nayyeri, B. Xiong, **Y. He**, E. Kharlamov, and S. Staab, "Predictive multiplicity of knowledge graph embeddings in link prediction"

SKILLS LIST

- **Research Areas**
Knowledge Graph Reasoning • Knowledge Graph–LLM Integration • Neuro-Symbolic AI • Agentic AI • RAG • Complex Logical Query Answering • Knowledge Representation and Reasoning • Information Retrieval
- **Knowledge Graph Methods**
Knowledge Graph Embeddings • Graph Neural Networks • Ontology Learning • Description Logics • Reasoning over Incomplete Knowledge Graphs • Multi-hop Question Answering • Hybrid Structured–Textual Reasoning
- **Machine Learning and Generative AI**
Large Language Models • Foundation Models • Graph Attention Networks • Reliable and Interpretable ML • Hallucination Mitigation • Knowledge Extraction from LLMs • Multimodal Reasoning
- **Technologies**
Python • PyTorch • PyTorch Geometric (PyG) • RDF • SPARQL • OWL • Neo4j/Cypher • TypeDB/TypeQL • Hugging Face Transformers • Git • Linux

Research Leadership & Service

- Supervised Bachelor and Master research projects on robust KG embeddings and KG-based RAG.
- Served as program chair for AAAI 2026, reviewed for ACL, ISWC, WWW, ESWC, and COLING.
- Collaborated across university and industrial research teams at SAP and Bosch.