

# YIFU QIU

[yifu.qiu@ed.ac.uk](mailto:yifu.qiu@ed.ac.uk) | [yfqiu.netlify.app](https://yfqiu.netlify.app) | [Google Scholar](#) | +44 07873801958

## PROFESSIONAL SUMMARY

Final-year PhD Researcher at University of Edinburgh/Cambridge (via ELLIS) specializing in **Foundation Model (LLMs & VLMs)**, **Post-Training, hallucination detection and mitigation**, and understanding their **internalized world model**, with first-author publications in ICLR, NeurIPS, ACL, EMNLP, and NAACL. Proven industrial research experience/collaboration at **Meta FAIR**, **Apple AIML** (with **Apple AIML PhD Fellowship**), **Nvidia Research**, and **Baidu NLP**, contributing to the large-scale pre/post-training for an omnimodal models (7B+ params), long-context LLMs' post-training for retrieval and reasoning, and hallucination mitigation. Expert in scaling distributed training pipelines and designing novel alignment algorithms (SFT/RL).

## EDUCATION

<b>University of Edinburgh &amp; University of Cambridge</b> <i>PhD in Natural Language Processing (ELLIS Program)</i>	Oct 2022 – Jun 2026 (Expected) United Kingdom
– <b>Topic:</b> Internalized World Models in Foundation Models (Advisors: Shay Cohen, Anna Korhonen, Edoardo Ponti). – Fully funded by <b>Apple AIML PhD Fellowship</b> and Informatics Global PhD Scholarship. – Selected for the ELLIS PhD program (< 5% acceptance rate).	
<b>University of Edinburgh</b> <i>MSc in Cognitive Science (NLP Specialism) – Distinction</i>	Sep 2020 – Aug 2021 United Kingdom
<b>BNU-HKBU United International College</b> <i>BSc in Computer Science – 1st Class Honours</i>	Sep 2016 – Aug 2020 China

## RESEARCH EXPERIENCE

<b>Research Scientist Intern</b> <i>Meta FAIR (Fundamental AI Research)</i>	June 2025 – Sep 2025 Paris, France
Project: <b>Omnimodal Foundation Models (Large Concept Model)</b>   Host: <i>Dr. Holger Schwenk</i> – Developed vSONAR, a vision-language embedding space supporting 200+ text and 37 speech languages. – Introduced vLCM, a scalable post-training pipeline for the 7B Large Concept Model (LCM), utilizing <b>1M instruction tuning</b> samples and 200K interleaved multimodal Chain-of-Thought data. – Achieved SOTA performance on multilingual captioning and QA, outperforming existing baselines in <b>61 out of 62 languages</b> . – Paper: <i>Unified Vision-Language Modeling via Concept Space Alignment</i> (ICLR 2026).	
<b>Research Scientist Intern</b> <i>Apple AIML</i>	June 2024 – Sep 2024 Seattle, WA
Project: <b>Long-Context Reasoning &amp; Inference Optimization</b>   Hosts: <i>Dr. Benjamin Han, Dr. Yizhe Zhang</i> – Proposed <b>ICR<sup>2</sup></b> , a benchmark for evaluating In-Context Retrieval and Reasoning in realistic scenarios. – Engineered <b>Retrieval-Attention-Probing</b> , an inference-time alignment technique using attention heads to de-noise long contexts during decoding. – Demonstrated +17% Exact Match gain on LOFT and +13% on ICR <sup>2</sup> using Mistral-7B, outperforming GPT-4-Turbo on specific tasks. – Paper: <i>Eliciting In-context Retrieval and Reasoning for Long-Context Language Models</i> (ACL 2025).	
<b>Research Scientist Intern</b> <i>Apple AIML</i>	June 2023 – Sep 2023 Seattle, WA
Project: <b>Trustworthy Generation &amp; Hallucination Mitigation</b>   Host: <i>Dr. Benjamin Han</i> – Developed <b>TWEAK</b> , a decoding-only hypothesis verification method to reduce hallucinations in knowledge-to-text generation without model retraining. – Created <b>FATE</b> (Fact-Aligned Textual Entailment), a dataset for training task-specific verifiers. – Improved faithfulness metrics (FactKB) by <b>7.17 points</b> in out-of-distribution settings with negligible latency cost. – Paper: <i>Think While You Write: Hypothesis Verification Promotes Faithful Knowledge-to-text Generation</i> (NAACL 2024).	
<b>Research Scientist Intern</b> <i>Baidu, Inc. (NLP Department)</i>	Nov 2021 – Aug 2022 Beijing, China
Project: <b>Dense Retrieval at Scale</b>   Hosts: <i>Dr. Hongyu Li, Dr. Jing Liu</i> – Constructed <b>DuReader-Retrieval</b> , the largest Chinese passage retrieval benchmark derived from real-world search engine logs.	

- Implemented diversity-enhanced dense retrieval training strategies using Natural Language Inference supervision.
- Paper: *DuReader-Retrieval: A Large-scale Chinese Benchmark...* (EMNLP 2022).

## PUBLICATION LIST

---

### **Self-Improving World Modelling with Latent Actions** (Collaboration with Nvidia Research)

Yifu Qiu, Zheng Zhao, Waylon Li, Yftah Ziser, Anna Korhonen, Shay Cohen, Edoardo Ponti  
*Under Review, ICML 2026*

### **Unified Vision-Language Modeling via Concept Space Alignment**

Yifu Qiu, Paul-Ambroise Duquenne, Holger Schwenk  
*ICLR 2026*

### **Bootstrapping Action-Grounded Visual Dynamics in Unified Vision-Language Models** (Collaboration with Nvidia Research)

Yifu Qiu, Yftah Ziser, Anna Korhonen, Shay Cohen, Edoardo Ponti  
*Under Review, ACL 2026*

### **Eliciting In-context Retrieval and Reasoning for Long-Context Language Models**

Yifu Qiu, Varun Embar, Shay Cohen, Yizhe Zhang, Navdeep Jaitly, Benjamin Han  
*ACL 2025*

### **Spectral Editing of Activations for Large Language Model Alignment**

Yifu Qiu, Zheng Zhao, Anna Korhonen, Edoardo Ponti, Shay Cohen  
*NeurIPS 2024*

### **Are Large Language Models Temporally Grounded?**

Yifu Qiu, Zheng Zhao, Anna Korhonen, Edoardo Ponti, Shay Cohen  
*NAACL 2024*

### **Think While You Write: Hypothesis Verification Promotes Faithful Knowledge-to-text Generation**

Yifu Qiu, Varun Embar, Shay Cohen, Benjamin Han  
*NAACL 2024*

### **Detecting and Mitigating Hallucinations in Multilingual Summarisation**

Yifu Qiu, Anna Korhonen, Edoardo Ponti, Shay Cohen  
*EMNLP 2023*

### **Abstractive Summarization Guided by Latent Hierarchical Document Structure**

Yifu Qiu, Shay Cohen  
*EMNLP 2022*

### **DuReader-Retrieval: A Large-scale Chinese Benchmark for Dense Information Retrieval from Real-World Applications**

Yifu Qiu, Hongyu Li, Yingqi Qu, Ying Chen, Qiaoqiao She, Jing Liu, Hua Wu, Haifeng Wang  
*EMNLP 2022*

## COLLABORATION PUBLICATION LIST

---

### **Lost in Space? Vision-Language Models Struggle with Relative Camera Pose Estimation** (Collaboration with Nvidia Research)

Ken Deng, Yifu Qiu, Yoni Kasten, Shay Cohen, Yftah Ziser  
*Under Review, ACL 2026*

### **Iterative Multilingual Spectral Attribute Erasure**

Shun Shao, Zheng Zhao, Yifu Qiu, Yftah Ziser, Shay Cohen  
*EMNLP 2025*

### **What Is That Talk About? A Video-to-Text Summarization Dataset for Scientific Presentations**

Dongqi Liu, Chenxi Whitehouse, Xi Yu, Louis Mahon, Rohit Saxena, Zheng Zhao, Yifu Qiu, Mirella Lapata, Vera Demberg  
*ACL 2025*

### **Multiple Instance Learning with Coarse-to-Fine Self-Distillation**

Shuyang Wu, Yifu Qiu, Ines Nearchou, Sandrine Prost, Jonathan Fallowfield, Hakan Bilen, Timothy Kendall  
*Under Review by TMI*

### **EEE-QA: Exploring Effective and Efficient Question-Answer Representations**

Zhanghao Hu, Yijun Yang, Junjie Xu, Yifu Qiu, Pinzhen Chen  
*LREC-COLING 2024*

## TECHNICAL SKILLS

---

**Core Research Area:** LLM/VLM Post-Training (SFT, RLHF), Multimodality, Reasoning, AI Safety (truthfulness and fairness)

**Deep Learning:** PyTorch, HuggingFace Transformers, Diffusers

**Scaling and Infrastructure:** DeepSpeed, FSDP, Slurm, Distributed Data Parallel (DDP)

**Languages:** Python, C++, SQL, Bash

## AWARDS & HONORS

---

**Apple AIML PhD Fellowship** (2023) – Awarded to top PhD students in AI worldwide.

**ELLIS PhD Program Member** (2022) – Selected for excellence in European AI research.

**Informatics Global PhD Scholarship**, University of Edinburgh (2022).

**Lord Frederick Cavendish Award**, University of Cambridge (2022).

**Top 1% Scholarship**, BNU-HKBU UIC (2019, 2020).