

姓名：包启明      出生年月：1996.07.14      性别：男      籍贯：浙江杭州      民族：汉族  
微信：qiming\_bao      联系电话：17826833549      邮箱：[bqmbill714@gmail.com](mailto:bqmbill714@gmail.com)  
个人主页：<https://14h034160212.github.io/>      GitHub：<https://github.com/14H034160212>



## 自我评价

我的博士期间的研究方向包含自然语言处理和推理，在博士学校是2024 QS世界排名前100的学校。Google Scholar总引用超过350次，总论文16篇，H-index: 9，单篇论文引用超过100次的1篇，GitHub星标超过90次1项，代表性工作有1篇一作ACL Findings，1篇一作AAAI Proceeding，3篇一作IJCAI/ICLR的大模型和AGI的专题研讨会，1篇一作周志华老师背书的逻辑推理新会IJCLR。其中最具有代表性的工作是提出了AMR-LDA Prompt Augmentation的方法结合GPT-4在ReClor逻辑推理排行榜上排名第一（第一个在ReClor隐藏测试集上超过90%准确率的团队）。另外还有共同一作ACL Findings, IJCAI, AAAI Proceeding, ICLR（通讯作者）。另外我有超过五年国内外AI研究/开发工作和项目经验。并且我也在微软亚研院，英国三星AI研究中心，墨尔本大学NLP组，浙大NLP组和中科院自动化所，美国麻省大学阿默斯特校区(UMass)，宾州州立大学(PSU)进行自然语言处理和逻辑推理的特邀报告。目前有一个澳门城市大学助理教授的offer。

## 教育背景

2020年2月-2025年9月 新西兰奥克兰大学 计算机科学 博士（全额奖学金） 导师：Michael Witbrock 教授  
2018年7月-2019年9月 新西兰奥克兰大学 计算机科学 一等荣誉学士 导师：Jiamou Liu 副教授和 Paul Denny 教授  
2014年9月-2018年6月 中国计量大学 计算机科学与技术 本科 导师：陈晓竹教授

## 部分项目与工作经历

2020年2月-2025年9月 Strong AI Lab, NAOInstitute, 新西兰奥克兰大学  
大语言模型和自然语言推理相关的工作（博士主方向）

- 博士课题项目：Strong AI Lab 项目（资助号 5000675），由新西兰高等教育委员会（Tertiary Education Commission）通过创业研究基金（Entrepreneurial Research Funding）资助，总金额为 960 万新西兰元。我在该项目中主要负责逻辑推理研究方向。
- 开发基于 LLM 的迭代增强框架，来生成 Learnersourced 平台多项选择题的解释。该框架通过解释生成和解释评估模块的迭代交互，以提高生成解释的质量。该工作已经被 AGI@ICLR 2024 和 AAAI Proceeding 2025 接受。[\[论文\]](#)[\[代码\]](#)
- 我们的模型 AMR-LDA 在逻辑推理阅读理解竞赛 ReClor 上取得第 1 名的成绩，相关工作已经被 ACL Findings 2024 和 LLM@IJCAI 2023 接受。该竞赛由新加坡国立大学主办，该竞赛收录了美国 LSAT 法律考试的题目，是目前自然语言推理任务中难度最高的两个竞赛之一。[\[论文\]](#)[\[代码\]](#)[\[模型\]](#)[\[排行榜\]](#)
- 我们在分布外（Out-of-Distribution）逻辑推理任务上评估了生成式和判别式大型语言模型。虽然这些模型在标准基准任务中表现良好，但输入的微小变化会导致显著的性能下降，凸显了它们在推理能力方面的不足。我们的论文已被 LLM@IJCAI'23 [\[论文\]](#) 和 IJCAI 2024 [\[论文\]](#) 接收（已被引用超过 100 次），相关的[\[代码\]](#)已在 GitHub 上公开。
- 为了解决现有 Multi-step reasoning 数据集存在的 Depth Imbalance 的问题，我们设计并开发了一个多跳推理自然语言数据集 PARARULE-Plus。基于 DeepLogic 和 Gate Attention 开发了 IMA-GloVe-GA 模型，模型在 OOD Multi-step reasoning 任务上的效果超过了其他基于 RNN 和预训练 Transformer 的 baseline 模型。相关工作在 IJCLR-NeSy 2022 上发表。我们在 IMA-GloVe-GA 论文中采用的门控注意力（Gated Attention）机制已被引用，并对 Qwen3 模型[\[论文\]](#)的提出提供了灵感。[\[论文\]](#)[\[代码和数据\]](#)[\[演讲视频\]](#)
- 基于 PARARULE-Plus 数据集研发溯因推理数据集，使神经网络具备一定的逻辑推理能力和推理的可解释性。相关工作在 ACL Findings 2022 上发表。PARARULE-Plus 和 AbductionRules 这两个数据集目前已被 [LogiTorch.ai](#), [ReasoningNLP](#), [Prompt4ReasoningPapers](#), [OpenAI/Evals](#), [A Survey on Evaluation of Large Language Models](#) 和 [Reasoning Language Models: A Blueprint](#) 收录。[\[论文\]](#)[\[代码和数据\]](#)

2022年7月-至今 新西兰人工智能软件公司 Xtracta（新西兰白名单公司）  
人工智能研究员/工程师

- 使用 PEFT Adapter 对多模态大模型 Qwen2-VL 和 InternVL2 的 Qwen2 子模型训练，并结合持续训练，可以在一块 A6000 GPU 上对 1B 参数的 InternVL2 多模态大模型进行训练。

- 调查并实施更长的注意力机制以增加多模态预训练 Transformer 模型(包括 LayoutLMv3 和 ERNIE-LayoutX)的有效序列长度来帮助模型对长文本的信息提取任务。整合 Global Attention Mask 来加强模型对文本向量表示和视觉向量表示的学习。
- 通过应用滑动窗口技术和 Longformer 的全局注意力掩码, 将最大序列长度从 512 扩展到 4096, 在不显著增加 GPU 内存使用的情况下, 帮助 LayoutLMv3 和 ERNIE-LayoutX 在 XFUND、FUNSD 和其他公司内部数据集上在 Token Classification 和 Relation Extraction 任务上获得更高的 F1 分数。
- 复现微软没有开源的 LayoutLMv3 多模态预训练代码, 包括 mask language modeling、mask image modeling、word-patch alignment。
- 成功申请新西兰创新机构 Callaghan Innovation 的研发税收激励 (RDTI) 资助, 分别针对 2022 年和 2023 年, 每年提供相当于研发支出 15% 的税收抵免。
- 将 Flash-Attention 2 集成到 Self-Attention 中, 可以帮助 ERNIE-LayoutX 在 FP16 下将最大训练 GPU 内存使用量减少多达 50%。
- 应用仿射变换进行数据增强, 以训练模型并提高文档提取过程中的对齐鲁棒性(line alignment issue)的问题。
- 通过使用 PEFT 适配器, Flash-Attention 2 和 int4 量化来持续训练 Qwen2-VL-7B, 并使 Qwen2-VL-7B 在单个 A4090 GPU (24GB GPU 内存内) 上进行训练。

2019 年 11 月 -2020 年 2 月

北京大学高等信息技术研究院 (杭州)

NLP 研发实习工程师

- 针对会议机器人自然语言处理技术的研究与开发, 包括自动摘要提取, 文本分割和主题预测。
- 机器人相关技术的调查和标注文档。
- 构建一个封装良好的 API, 以基于摘要提取, 文本分段和主题预测来实现会议记录文档整理。

2018 年 11 月 -2019 年 4 月

新西兰医疗公司 Precision Driven Health

暑期研究项目

- 基于深度学习和知识图谱, 提出了一种基于 word attention 的医疗文本相似度计算模型 HBAM。
- 我们的工作被澳洲计算机年会 ACSW 2020 收录, 目前已超过 80 次引用, 代码超过 90 次星标保存。[\[论文\]](#)[\[代码\]](#)[\[新闻\]](#)[\[演讲\]](#)

已发表或接收的文章

- [Qiming Bao](#), Juho Leinonen, Alex Yuxuan Peng, Wanjun Zhong, Tim Pistotti, Alice Huang, Paul Denny, Michael Witbrock, Jiamou Liu. Exploring Iterative Enhancement for Improving Learnersourced Multiple-Choice Question Explanations with Large Language Models, **Proceedings of the AAAI Conference on Artificial Intelligence (2025) (Core Rank A\*, CCF A 类顶会, AAAI Proceeding)**
- [Qiming Bao](#), Alex Peng, Zhenyun Deng, Wanjun Zhong, Gaël Gendron, Neşet Tan, Nathan Young, Yang Chen, Yonghua Zhu, Michael Witbrock, Jiamou Liu. Abstract Meaning Representation-Based Logic-Driven Data Augmentation for Logical Reasoning., **The Findings of ACL (2024) (Core Rank A\*, CCF A 类顶会)**
- [Qiming Bao](#), Juho Leinonen, Alex Yuxuan Peng, Wanjun Zhong, Tim Pistotti, Alice Huang, Paul Denny, Michael Witbrock, Jiamou Liu. Exploring Iterative Enhancement for Improving Learnersourced Multiple-Choice Question Explanations with Large Language Models, **AGI@ICLR (2024) (Core Rank A\*顶会的研讨会)**
- [Qiming Bao](#), Gaël Gendron, Alex Peng, Neset Tan, Michael Witbrock, Jiamou Liu. Assessing and Enhancing the Robustness of Large Language Models with Task Structure Variations for Logical Reasoning., **ICONIP (2024) (Core Rank B, CCF C, 新西兰奥克兰举办, EI 检索)**
- [Qiming Bao](#), Gaël Gendron, Alex Peng, Neset Tan, Michael Witbrock, Jiamou Liu. A Systematic Evaluation of Large Language Models on Out-of-Distribution Logical Reasoning Tasks., **LLM@IJCAI (2023) (Core Rank A\*, CCF A 类顶会的大语言模型研讨会)**
- [Qiming Bao](#), Alex Peng, Zhenyun Deng, Wanjun Zhong, Gaël Gendron, Neşet Tan, Nathan Young, Yang Chen, Yonghua Zhu, Michael Witbrock, Jiamou Liu. Enhancing Logical Reasoning of Large Language Models through Logic-Driven Data Augmentation., **LLM@IJCAI (2023) (Core Rank A\*, CCF A 类顶会的大语言模型研讨会)**
- [Qiming Bao](#), Alex Peng, Tim Hartill, Neset Tan, Zhenyun Deng, Michael Witbrock, Jiamou Liu. *Multi-Step Deductive*

*Reasoning Over Natural Language: An Empirical Study on Out-of-Distribution Generalisation*, **IJCLR-NeSy Long Paper (2022)** (周志华教授背书的学习和推理结合的专业领域会议, 2022 年第二届)

- Nathan Young, [Qiming Bao](#), Joshua Ljudo Bensemann, Michael J. Witbrock. *AbductionRules: Training Transformers to Explain Unexpected Inputs*, **The Findings of ACL (2022)** (Core Rank A\*, CCF A 类顶会)
- Gaël Gendron, [Qiming Bao](#), Michael Witbrock, Gillian Dobbie. Large Language Models Are Not Strong Abstract Reasoners, **IJCAI (2024)** (Core Rank A\*, CCF A 类顶会)
- Lin Ni, [Qiming Bao](#), Xiaoxuan Li, Qianqian Qi, Paul Denny, Jim Warren, Michael Witbrock, Jiamou Liu. *DeepQR: Neural-based Quality Ratings for Learnersourced Multiple-Choice Questions*, **Proceedings of the AAAI Conference on Artificial Intelligence (2022)** (Core Rank A\*, CCF A 类顶会, AAAI Proceeding)
- Qianqian Qi, [Qiming Bao](#), Alex Yuxuan Peng, Jiamou Liu, Michael Witbrock. A Dynamic Prompt-tuning Method for Data Augmentation with Associated Knowledge, **ICLR TinyPapers (2023)** (Core Rank A\*顶会)
- Gaël Gendron, [Qiming Bao](#), Michael Witbrock, Gillian Dobbie. Large Language Models Are Not Strong Abstract Reasoners Yet, **AGI@ICLR (2024)** (Core Rank A\*顶会的研讨会)
- [Qiming Bao](#), Lin Ni, Jiamou Liu. *HHH: An Online Medical Chatbot System based on Knowledge Graph and Hierarchical Bi-Directional Attention*, **ACSW (2020)** (EI 检索)
- Zhongsheng Wang, Jiamou Liu, [Qiming Bao](#), Hongfei Rong, Jingfeng Zhang. ChatLogic: Integrating Logic Programming with Large Language Models for Multi-step Reasoning, **NucLeaR@AAAI (2024)** (Core Rank A\*, CCF A 类顶会的研讨会)
- Neset Ozkan TAN, Trung Nguyen, Josh Bensemann, Alex Peng, [Qiming Bao](#), Yang Chen, Mark Gahegan, Michael Witbrock. Multi2Claim: Generating Scientific Claims from Multi-Choice Questions for Scientific Fact-Checking, **EACL (2023)** (Core Rank A 类顶会)
- Neset Tan, Alex Peng, Joshua Bensemann, [Qiming Bao](#), Tim Hartill, Mark Gahegan, Michael Witbrock. *Input-length-shortening and text generation via attention values*, **AAAI-EMC<sup>2</sup> (2023)** (Core Rank A\*, CCF A 类顶会的研讨会)

**个人荣誉** (在大学本科期间, 每学期都评为校三好学生, 并六次获得一等奖学金, 综合排名: 1/120)

2018 年 2 月	<a href="#">国际大学生数学建模竞赛二等奖</a>	数学建模协会
2018 年 6 月	<a href="#">浙江省优秀毕业生</a>	浙江省教育厅
2016 年 12 月	浙江省多媒体竞赛移动应用类三等奖	浙江省教育厅
2020 年 12 月	新西兰硕博会篮球赛最有价值球员和三分比赛冠军	新西兰硕博会
2022 年 12 月	<a href="#">计算机系优秀博士学长(Outstanding PhD Mentor)</a>	新西兰奥克兰大学
2023 年 4 月	企业研发税收激励 (RDTI) 资助	新西兰创新机构 Callaghan Innovation
2024 年 4 月	<a href="#">中国计量大学澳新校友会副会长</a>	中国计量大学
2024 年 6 月	<a href="#">新西兰奥克兰大学优秀博士生</a>	新西兰奥克兰大学
2025 年 4 月	<a href="#">DAAD AINeT fellow 2025 on Natural Language Processing</a>	德国学术交流服务处 German Academic Exchange Service

## 学术会议特邀报告和访问学者(Invited Talk/Visiting Scholar)

- 微软亚洲研究院 特邀报告 2022 邀请人: 陈蓓博士 ([邀请函](#)) ([演讲稿](#)) ([演讲回放](#))
- 英国三星 AI 研究中心 特邀报告 2022 邀请人: Cristina Cornelio 博士 ([邀请函](#)) ([演讲稿](#)) ([演讲回放](#))
- IEEE Vehicular Technology Society (VTS) New Zealand North Chapter and IEEE New Zealand North Section SIGHT Group 特邀报告 2022 邀请人: William Liu 教授 ([邀请函](#)) ([演讲稿](#)) ([演讲回放](#))
- 浙江大学 ZJU-NLP Group 访问学者 2023 邀请人: 陈华均教授和张宁豫教授
- 墨尔本大学 NLP Group 特邀报告 2023 邀请人: Ming-Bin Chen 和 Jey Han Lau 教授 ([邀请函](#)) ([演讲稿](#))
- 中科院自动化所 特邀报告 2023 ([邀请函](#)) ([演讲稿](#))
- 深圳北理莫斯科大学 特邀报告 2024 ([邀请函](#)) ([演讲稿](#))
- 美国麻省大学阿默斯特校区(UMass) 特邀报告 2024 邀请人: Andrew Lan 教授 ([邀请函](#)) ([演讲稿](#))
- 美国宾州州立大学(Penn State University) 特邀报告 2024 ([邀请函](#)) ([演讲稿](#)) ([演讲回放](#))
- 北京大学和清华大学联合 Logic AI Seminar 特邀报告 2025 邀请人: 刘奋荣教授和李昊轩教授 ([邀请函](#))
- 德国马普所 Max Planck Institute For Software Systems 特邀报告 2025 邀请人: Adish Singla 教授 ([邀请函](#))
- 德国慕尼黑工业大学 TUM 特邀报告 2025 邀请人: Stefan Fuchs 博士和 André Borrmann 教授