

# 张鹏宇

## 个人信息

<https://pengyu-zhang.github.io/>

p.zhang@uva.nl

15801349669

博士三年级在读

研究方向：多模态知识图谱、知识获取、图神经网络、链接预测

## 教育经历

- 2022.10 - 至今 阿姆斯特丹大学 (**University of Amsterdam**)，阿姆斯特丹，荷兰  
信息学研究所 (Informatics Institute) 智能数据工程实验室 (Intelligent Data Engineering Lab)  
博士研究生  
导师：Prof. Dr. Paul Groth ([link](#))
- 2019.09 - 2022.07 北京工业大学，北京，中国  
控制工程 硕士; GPA: 3.49 / 4; 排名: 7/81  
导师：Prof. Dr. 张勇([link](#))

## 项目经历

- 2025.01 - 至今 **MMKG**: 将“时间”作为独立模态的多模态知识图谱建模  
**问题**: 在多模态知识图谱中，多数实体存在“高度相似、难以区分”的情况，导致表示区分度不够、预测表现受限。  
**方法**: 时间独立模态化：将时间作为与文本/图像同等地位的模态；用时间编码模块对时间戳建模；多时间戳选择与聚合，用注意力分配权重并聚合为统一时间向量，兼顾信息完整与鲁棒性。  
**成效**: 在 3 个公共基准上取得 SOTA，对“文本/图像高度相似”的实体（如历史人物 vs 同名电影）显著降低混淆。
- 2024.03 - 2024.12 **LLM-MMKG**: 用 LLM 增强多模态知识图谱  
**问题**: 多模态知识图谱中大量图片语义含糊（如标志、抽象画）或噪声多，人工筛选成本高且有偏见。  
**方法**: 自动检索实体相关图片与元数据（含时间戳、描述），扩展视觉覆盖面；用预训练视觉-语言模型把“语义稀疏/过度复杂”的图片转成可用文本描述；用 LLM 做多描述汇总与降噪，保留与实体与任务最相关的信息，作为“语言桥梁”对齐图像与实体语义。  
**成效**: 在 MKG-W、MKG-Y、DB15K 三个公共数据集上构建自动化流水线；清洗、对齐并发布增强版数据集（扩展图片与配套文本描述），配套代码与说明文档。
- 2023.08 - 2024.01 **CYCLE**: 面向时间演化的实体链接与公平性提升  
**问题**: 知识图谱随时间演化（新实体出现、描述与关系更新），模型易出现“时间退化”，尤其伤害低度节点表现。  
**方法**: 把图结构信息与文本信息结合，设计基于时间信号的动态对比机制。利用时间标注引导正负样本的对比学习；联合图嵌入与文本表示以降低结构不公平。  
**成效**: 缓解时间退化趋势，低度节点召回更稳定；方案可无缝接入现有实体链接流水线。
- 2022.09 - 2023.08 **TIGER**: 时间稳健的实体可区分性建模  
**问题**: 知识图谱更新后，实体链接模型在新时间点上性能下滑。  
**方法**: 面向时间切片对齐训练；对关系上下文与描述文本做协同编码，减少时间漂移带来的语义偏差。

		成效：在跨时间评测上更稳健，迁移到新图谱版本时性能下降更小。
2021.07 - 2022.07	<b>RUNCL</b>	面向幂律度分布的节点分类提升 问题：很多 GNN 数据呈幂律分布，少数高连接节点主导学习，低度节点效果差、训练不稳。 方法：通过图生成获得多样邻域视图，并做最优图选择；配合稳健的特征学习流程，抑制高低度不平衡。 成效：节点分类更稳；低度节点的准确率与一致性明显提升。
2020.07 - 2021.06	<b>MVMA-GCN</b>	多视图多层次注意力的半监督节点分类 问题：真实网络往往有多种节点与多种关系，单一图难以完整表达，信息整合困难。 方法：为不同关系构建独立视图，使用注意力做跨视图融合；用 HSIC (Hilbert–Schmidt 独立性准则) 区分与约束关系差异，避免信息混淆。 成效：得到更干净的节点表征，在半监督节点分类任务上稳定增益。
2019.11 - 2021.04	科研论文作者消歧	科研论文作者消歧：科研论文作者消歧可视分析系统 问题：服务科研管理的作者重名消歧。 成效：构建了可视化系统，包含界面以及后台算法，把复杂网络信息变得可探索、可解释，辅助人工高效完成消歧决策。

## 相关论文

- 在审(双盲) Pengyu Zhang, Klim Zaporojets, Congfeng Cao, Jia-Hong Huang and Paul Groth.
- 在审(双盲) Pengyu Zhang, Klim Zaporojets, Jie Liu, Jia-Hong Huang and Paul Groth
- 在审(双盲) Fina Polat, Pengyu Zhang, Daniel Daza, Klim Zaporojets and Paul Groth
- 2025.09 Mengran Li, Pengyu Zhang, Wenbin Xing, Yijia Zheng, Klim Zaporojets, Junzhou Chen, Ronghui Zhang, Yong Zhang, Siyuan Gong, Jia Hu, Xiaolei Ma, Zhiyuan Liu, Paul Groth, Marcel Worring. [A survey of large language models for data challenges in graphs](#). *Expert Systems with Applications*.
- 2024.11 James Nevin, Pengyu Zhang, Dimitar Dimitrov, Michael Lees, Paul Groth, Stefan Dietze. [Understanding the Impact of Entity Linking on the Topology of Entity Co-occurrence Networks for Social Media Analysis](#). *International Conference on Knowledge Engineering and Knowledge Management*. Cham: Springer Nature Switzerland, 2024.
- 2024.07 Pengyu Zhang, Congfeng Cao, Klim Zaporojets, Paul Groth. [CYCLE: Cross-Year Contrastive Learning in Entity-Linking](#). *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. 2024.
- 2024.06 Pengyu Zhang, Congfeng Cao, Paul Groth. [TIGER: Temporally Improved Graph Entity Linker](#). *ECAI 2024*. IOS Press, 2024. 3733-3740.
- 2024.01 Pengyu Zhang, Yong Zhang, Xinglin Piao, Yongliang Sun, Baocai Yin. [Relationship updating network with contrastive learning](#). *Physica A: Statistical Mechanics and its Applications*. 646 (2024): 129874.
- 2023.09 Pengyu Zhang, Yong Zhang, Jingcheng Wang, Baocai Yin. [MVMA-GCN: Multi-view Multi-layer Attention Graph Convolutional Networks](#). *Engineering Applications of Artificial Intelligence*. 126 (2023): 106717.
- 2023.04 Pengyu Zhang, Yong Zhang, Yanjie Cui and Baocai Yin. [Visual Analysis for Name Disambiguation of Academic Papers](#). (in Chinese) *Journal of Computer-Aided Design and Computer Graphics*. 2022, 34(11): 1659-1672.
- 2022.08 Pengyu Zhang, Yong Zhang, Yanjie Cui and Baocai Yin. [Visualization Analysis Method for Name Disambiguation of Academic Papers](#). (in Chinese) *CAD & CG 2021*.
- 2021.09 Xin Zheng, Pengyu Zhang, Yanjie Cui, Rong Du and Yong Zhang. [Dual-Channel Heterogeneous Graph Network for Author Name Disambiguation](#). *Information* 12.9 (2021): 383.