

# 学术报告

受中国矿业大学信息与控制工程学院邀请，中国科学院计算所许倩倩教授在我校举行线上学术报告。欢迎广大师生踊跃参加！

**报告题目：**跨媒体知识图谱构建与学习

**报告时间：**2023.10.20 日下午 14:00

**腾讯会议：**704-152-547

**报告人简介：**许倩倩，中国科学院计算所，研究员，博士生导师，基金委优秀青年基金获得者。CCF 杰出会员，IEEE/CSIG/高级会员，担任 CSIG 青工委副秘书长、CSIG 多媒体专委会副秘书长、CAAI 深度学习专委会副秘书长。研究领域为数据挖掘和机器学习，共发表 CCF-A 类论文 80 余篇（其中 TPAMI 15 篇）。先后获得：吴文俊人工智能自然科学一等奖、吴文俊人工智能科技进步二等奖、茅以升北京青年科技奖、中国图象图形学学会石青云女科学家奖、吴文俊人工智能优秀青年奖、中国人工智能学会最佳青年科技成果奖、ACM 中国



SIGMM 新星奖、北京图象图形学学会最美女科技工作者、中国人工智能学会优秀博士学位论文、中科院百篇优秀博士学位论文等奖励，并入选首份 AI 华人女性青年学者榜单。担任国际期刊 TMM、T-CSVT、ACM TOMM 和 Multimedia Systems 编委，CCF-A 类国际会议 ACM MM 领域主席，AAAI 和 IJCAI SPC。指导/协助指导的学生先后获得中国计算机学会优博、中国电子学会优博、中国科学院百篇优博（2 人次）、微软学者奖学金、中国科学院院长特别奖、百度奖学金全球 20 强等荣誉。

**报告摘要：**人工智能主要包含数据驱动和知识驱动两大范式。传统机器学习方法主要依赖数据驱动范式，依赖海量数据，外推能力差，通用性与泛化能力有限。与之相对，知识驱动范式可有效集成仅人类所能理解的复杂领域知识，可解释性强，具有更好的通用性，可与数据驱动范式优势互补。近年来，知识图谱因其高效的组织形式成为知识的主要载体，在诸多行业取得成功应用，使得知识驱动范式逐渐受到业界重视。同时，国家标准委、网信办等五部分联合印发的《国家新一代人工智能标准体系建设指南》中，将机器学习和知识图谱同时列为人工智能的关键通用技术。由此可见，数据知识联合驱动的范式势将成为新一代人工智能的显著特征。本次报告将聚焦于社交媒体、网络空间安全两个重大领域，介绍报告人在跨媒体知识图谱构建及学习方面的最新进展。针对图谱构建，将从数据采集、数据存储、图谱结构设计、系统接口服务、知识图谱可视化等多个方面分享跨媒体知识图谱构建的基本思路和实践经验。针对图谱构建过程中遇到的数据标注不可信难题，介绍所提出的基于 HodgeRank 的异常样本发现框架；该框架从几何拓扑视角定量挖掘异常标注样本的数学特征，从而保证图谱中知识的准确性。同时，在图谱下游应用中，传统模型以准确率作为优化准则，仅考虑单一阈值下性能，而实际场景中阈值往往随需求动态变化，因而难以保证模型预测结果的可靠

性。针对该问题,将介绍所提出的基于 X-curve 的模型优化框架,该框架以 AUROC、AUPRC、AUTKC 等性能曲线为模型优化准则,综合考虑不同阈值下的模型性能,从而保证模型在复杂应用场景下的可靠性。

**欢迎广大师生参加!**

**中国矿业大学**

**信息与控制工程学院**

**2023 年 10 月 19 日**