

华中科技大学

研究生（参加校内外公开学术活动）报告

题目：AI 大模型时代下数据管理报告

学	号	<u>M202273662</u>		
姓	名	<u>金灿果</u>		
专	业	<u>计算机科学与技术</u>		
指	导	教	师	<u>左琼</u>
院（系、所）		<u>计算机学院</u>		

华中科技大学研究生院制

填表注意事项

- 一、本表适用于攻读硕士学位研究生选题报告、学术报告，攻读专业硕士学位研究生实践环节报告，攻读博士学位研究生文献综述、选题报告、论文中期进展报告、学术报告等。
- 二、以上各报告内容及要求由相关院（系、所）做具体要求。
- 三、以上各报告均须存入研究生个人学籍档案。
- 四、本表填写要求文句通顺、内容明确、字迹工整。

报告内容

2023 年 9 月 22 日，计算机学院计算机科学与技术专业建设 50 周年学术论坛——AI 大模型时代下的数据管理专场。

东北大学袁野教授分享了神经符号数据库的前沿应用并剖析其在数据库管理领域发展带来的深刻影响。其作为数据库的前沿应用，融合了人工智能分析、数据库查询与大数据计算技术，可高效处理结构化、半结构化及非结构化（如图像、视频、文本等）的异构数据。袁野教授对神经符号数据库在数据密集型领域的应用前景、技术挑战与解决思路均予以详细阐释。

在数字化进程加速的时代，数据成为数字化转型的关键支撑，仿若新能源，而数据管理系统则是数据驱动应用的根基。华东师范大学数据科学与工程学院院长钱卫宁教授深入探究数据管理系统的重要性，指出传统数据库系统难以契合新兴应用需求。数字化转型催生众多新兴应用，其数据处理需求远超传统数据库系统的能力范畴。尤其在人工智能应用中，模型计算需求剧增，传统数据库系统在满足此类需求时力不从心。例如深度学习模型训练与推理时，传统数据库架构难以在数据管理与模型计算间高效协同，形成应用性能瓶颈。由此，构建融合数据管理与模型计算的新型数据管理系统至关重要。该系统能打破传统隔阂，通过创新架构与算法优化，实现数据存储、管理与计算的深度融合，为新兴应用提供有力支撑。

在信息检索领域，清华大学张瑞教授阐述了大语言模型在该领域所蕴含的机遇以及需应对的挑战，并且详细介绍了多模态大模型和知识图谱调用等关键领域的最新进展，而这些进展对于提高回复正确率有着重要意义。着重探讨了当前研究的热点内容，涵盖多模态个性化预训练、迁徙学习算法以及自动进行知识图谱对齐的算法等。这些研究成果能够为知识图谱及信息检索相关的研究工作提供有力支持，助力其不断发展。

东北大学杨晓春教授深入探究了个性化推荐系统的重要性，着重指出该技术在多领域有着广泛应用。团队介绍了多个关键维度的研究成果，涵盖推荐效果提升、推荐可解释性研究以及不同推荐场景下的应用案例与探索。利用属性信息提高推荐的准确性，核心思想在于，对于一个 item 序列，在相关 attribute type 上一般会具有类似的 attribute value，显示偏好提取器将 attribute value 上具有显著共性的 attribute type 作为相关信息，参与 loss 和 uncertainty 的修正。强调个性化推荐算法在电子商务、社交媒体、新闻推荐、智能制造等诸多领域起着关键作用。其分享展现了该系统在不同领域的前沿研究与实际应用情况，能为相关研究及产业发展提供有益参考与启示，助力推动个性化推荐系统创新，提升其应用效果。

学习感悟

在数据学习和数据运用的过程中，要合理考虑数据价值，例如会议介绍的推荐算法提到序列与属性直接的关联性。复杂的数据往往需要挖掘出真正对用户有价值且精准的信息，这是数据应用的算法改进的关键突破口方向。

会议中数字化转型中的数据管理系统建设分析，启发学术领域深入探索数据管理理论创新，又助力产业界优化数据驱动应用实践。在推动数据管理系统技术革新的同时，还需要加速数字化转型步伐，引领数据管理领域迈向新发展阶段。数据管理并非一成不变，在未来的学习与实践中，应积极关注跨模态数据融合与知识图谱构建等技术，探索如何将其应用于实际的信息检索场景，如企业内部知识管理系统或专业领域的文献检索平台，以提升信息检索的效率与质量。

这些前沿学术也揭示了 AI 与数据管理领域的不同风险，如数据隐私与安全问题在数据管理、信息检索以及个性化推荐过程中都可能存在潜在风险；模型的可解释性虽然受到关注，但目前仍难以完全满足实际应用的需求；还有大规模数据处理的效率与成本控制等问题，都需要我们在未来的研究与实践中不断探索与攻克。

会议的学习创造了深入了解、学习该领域知识的良好契机，更为今后开展相关研究与实践活动提供了极具参考价值的有益指导。

研 究 生 签 字_____

指 导 教 师 签 字_____

院(系、所)领导签字_____

年 月 日