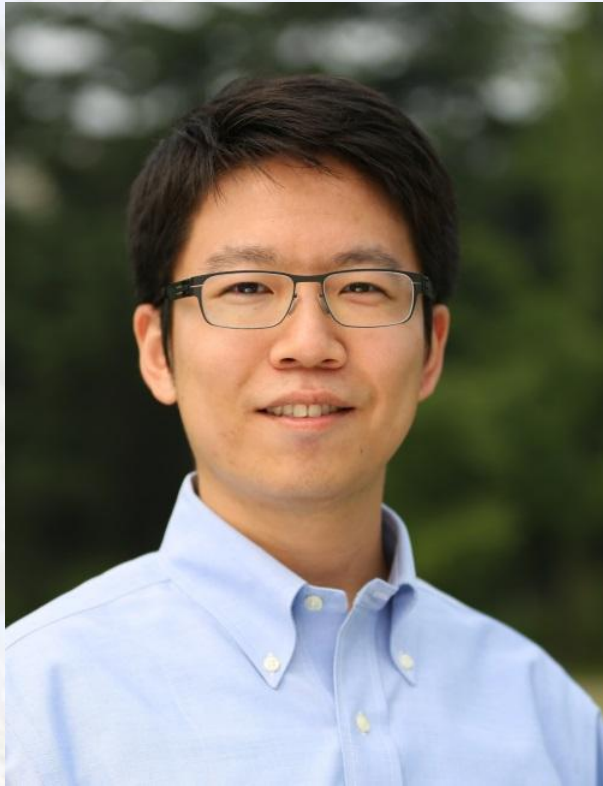


数据驱动的医学知识图谱构建

主讲人：俞声 (清华大学)

**TIME:**

14:00-15:30, April 4th, 2019
2019年4月4日14:00-15:30

VENUE:

Room 1303, Science Building No. 1
理科一号楼1303

CHAIR:

林伟 (北京大学)

摘要:

有效分析医疗大数据、实现基于知识的大规模人工智能建模和可靠临床决策支持离不开医学知识图谱的支持。医学知识图谱是以医学概念及其术语名称作为节点、以医学关系作为边的有向图，可以为电子病历分析中的术语识别、特征提取以及有效建模提供信息支持，是重要的医学信息基础设施。然而，由于传统的由专家收集整理信息的建设方式消耗巨大的资金与时间，至今世界范围内仍然没有一个较为完整的医学知识图谱。中文医学术语建设更是严重落后于发达国家，严重制约我国医疗大数据技术的发展。本报告介绍两方面的工作：1、利用图论和深度学习自动从电子病历数据中提取医学术语、建立医学术语库；2、利用深度学习模型和自然语言处理技术从维基百科和医学教科书自动提取医学关系图。

主讲人简介:

俞声博士的研究方向是医疗文本数据分析，主要研究内容包括自动术语识别、关系提取、表示学习等自然语言处理问题，以及大规模医学知识图谱构建、表型提取、临床决策支持等问题。俞声博士现任职清华大学统计学研究中心副教授、数据科学研究院RONG教授，归国前是电子病历驱动的基因组学研究先驱-i2b2美国国家生物医学计算中心的成员之一，在多项精准医学重点项目中承担研发任务。俞声博士将统计学与人工智能技术应用于医学信息领域，在电子病历文本数据分析领域取得了一系列突破性成果，每年在医学信息学顶级期刊JAMIA上发表论文，并获选Editor's Choice。他所开发的无监督学习技术使疾病表型识别算法开发速度从每年1-2个提高到每年超过1000个，并应用于Partners HealthCare Biobank、Veteran Affairs "Million Veteran Program"和eMERGE Network等美国国家级精准医学重点项目。