Utilizing Knowledge Graphs for Text-Centric Information Retrieval

* 来源：SIGIR’18, July 8-12, 2018, Ann Arbor, MI, USA
* 作者：Laura Dietz ；Alexander Kotov；Edgar Meij
* 简介：大型知识图谱（KG）和可扩展的实体链接技术是深入理解文本语义的强大工具。 本文重点关注知识图和实体链接如何最有效地用于以文本为中心的信息检索（IR）。
* 创新点及贡献：
  1. 实体链接：①识别可链接短语，即可以提及实体的文本片段。②检索每个这样的短语的候选集，当然必须考虑所提到的实体未包含在知识图中的可能性（所谓的NIL实体）。③根据提及的背景来区分提及提到的哪些候选实体。最佳实践方法，例如主题模型和单词嵌入。
  2. 实体检索：①在KG中查询通过其名称，属性或相关实体来引用目标实体，并以关键字或问题的形式表示②通过考虑知识图中的距离，可以使检索到的实体多样化③还可以使用基于语料库的伪相关性反馈方法来检索实体
  3. 以文本为中心的IR系统的三个核心角度：a）关键字匹配和平滑模型，b）使用伪相关反馈和查询日志查询扩展模型，以及c）用于多样化和冗余移除的组件

Weakly-supervised Contextualization of Knowledge Graph Facts

* 来源：SIGIR’18, July 8-12, 2018, Ann Arbor, MI, USA
* 作者：Nikos Voskarides；Edgar Meij ；Ridho Reinanda
* 简介：本文介绍了知识图谱事实情境化任务，并提出了NFCM（neural fact contextualization method），一种弱监督的方法来解决。
* 创新点及贡献：
  1. NFCM将手工制作的功能与使用深度学习自动识别的功能相结合。使用远程监督来通过使用KG中的高度重叠的大型实体标记文本语料库来增强训练数据的收集。
  2. ①远程监督是收集此任务的训练数据的有效手段，②NFCM明显优于此任务的几个启发式基线，③手工和自动学习的特征有助于NFCM的检索效果。

LogCanvas: Visualizing Search History Using Knowledge Graphs

* 来源：SIGIR’18, July 8-12, 2018, Ann Arbor, MI, USA
* 作者：Luyan Xu, Zeon Trevor Fernando, Xuan Zhou, and Wolfgang NejdlmNikos Voskarides；Edgar Meij ；Ridho Reinanda
* 简介：介绍了LogCanvas，一个基于图形的搜索历史可视化平台.。用户的搜索活动聚集到会话中。每个会话的查询都嵌入到知识图谱中，以帮助用户了解查询返回的结果以及它们的相关性。
* 创新点及贡献：
  1. 查询日志根据时间间隔聚类到会话中
  2. 对于每个会话的查询，使用Yahoo Fast Entity Linker 从搜索结果片段中提取相关概念和实体。
  3. 使用基于维基百科中的实体共现频率的方法来计算概念之间的相关性。

Improving Exploratory Search Experience through Hierarchical Knowledge Graphs

* 来源：SIGIR’17, August 7-11, 2017, Shinjuku, Tokyo, Japan
* 作者：Bahareh Sarrafzadeh；Edward Lank
* 简介：本文研究的主要目标是探索是否可以将知识图和层次结构的优势结合到一个数据结构中，以便可视化搜索结果。与分层树相比，分层图谱显着减少了文档读取时间，同时提供证据表明参与者通过分析交互日志来使用层次结构。
* 创新点及贡献：
  1. 使用解析算法自动生成知识图谱，使用动态阈值方法提取层次结构。评估层次知识图谱（HKG）作为低级实体关系和高级中心概念组合表示的效率。
  2. 使用混合方法方法评估这些HKG。定量数据认为HKG保留了知识图谱的透明度优势和层次结构的结构优势。定性数据与定量观察三角化，并提供对层次和网络可视化的优缺点的见解。

Entity Set Expansion via Knowledge Graphs

* 来源：SIGIR’17, August 7-11, 2017, Shinjuku, Tokyo, Japan
* 作者：Xiangling Zhang ；Jun Chen；Yueguo Chen
* 简介：本文提出了一种基于知识图中实体之间的语义相关性的分类模型（称为k-松弛的共同语义特征），以有效地评估实体与少数给定种子之间的相似性。 该模型旨在处理知识图谱的不完整性。
* 创新点及贡献：
  1. 使用解析算法自动生成知识图谱，使用动态阈值方法提取层次结构。评估层次知识图谱（HKG）作为低级实体关系和高级中心概念组合表示的效率。
  2. 使用混合方法方法评估这些HKG。定量数据认为HKG保留了知识图谱的透明度优势和层次结构的结构优势。定性数据与定量观察三角化，并提供对层次和网络可视化的优缺点的见解。

**Open-World Knowledge Graph Completion**

* 来源：The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)
* 作者：Baoxu Shi, Tim Weninger
* 简介：引入了一个名为ConMask的开放世界KGC模型。该模型学习实体名称及其文本描述部分的嵌入，以将看不见的实体连接到KG。
* 创新点及贡献：
  1. 为了减少噪声文本描述的存在，ConMask使用依赖于关系的内容屏蔽来提取相关片段，然后训练完全卷积神经网络以将提取的片段与KG中的实体融合。
  2. 由于标准KGC数据集中存在问题，我们还发布了两个用于KGC研发的新DBPedia数据集。

**Variational Reasoning for Question Answering with Knowledge Graph**

* 来源：The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)
* 作者：Yuyu Zhang；Hanjun Dai；Zornitsa Kozareva
* 简介：提出了一种新颖且统一的深度学习架构，以及端到端的变分学习算法，可以处理问题中的噪声，同时学习多跳推理。
* 创新点及贡献：
  1. 在文献中最近的基准数据集上实现了最先进的性能。
  2. 推导出一系列新的基准数据集，包括多跳推理的问题，神经翻译模型解释的问题以及人类语音中的问题。

**Complex Sequential Question Answering: Towards Learning to Converse Over Linked Question Answer Pairs with a Knowledge Graph**

* 来源：The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)
* 作者：Amrita Saha；Vardaan Pahuja；Mitesh M. Khapra
* 简介：介绍了复杂顺序问答（CSQA）的任务，它是对话和质量保证的最先进模型之间的交叉。
* 创新点及贡献：
  1. CSQA包含由链接QA对上的对话组成的大规模数据集。 该数据集包含200,000个对话，转速为1.6M，并通过手动密集的半自动化过程收集。最先进的性能。
  2. 这是第一个包含复杂问题的数据集，这些问题需要对包含数百万元组的大型知识图进行逻辑，定量和/或比较推理。