

Construction and Implementation of the Security Common Sense Guide Graph Based on Neo4j

Ning An¹, Haiying Zhao^{2*}

¹Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing

²School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing

Email: *zhaohaiying@bupt.edu.cn

Received: Jan. 4th, 2018; accepted: Jan. 18th, 2018; published: Jan. 25th, 2018

Abstract

In the Internet age, it is more and more important to carry out safety knowledge popularization through the network platform. The general purpose of safety common sense popularization lies in improving people's awareness of life safety and cultivating people's ability of self-protection and self-help in the event of an accident, but the current common sense of safety popularization cannot effectively reflect the relationship between various safety common senses. Based on the data of safety popularization net, the paper constructs the safety common sense map based on Neo4j, presents the data in a visualized form, builds a knowledge system with safety common sense as the core, and makes it convenient for users to query and learn safety common knowledge.

Keywords

Neo4j, Safety Common Sense, Common Sense Guide Graph

基于Neo4j的安全常识导图构建与实现

安 宁¹, 赵海英^{2*}

¹北京邮电大学网络技术研究院, 北京

²北京邮电大学数字媒体与设计艺术学院, 北京

Email: *zhaohaiying@bupt.edu.cn

收稿日期: 2018年1月4日; 录用日期: 2018年1月18日; 发布日期: 2018年1月25日

*通讯作者。

摘要

在互联网时代, 通过网络平台进行安全知识科普越来越重要。安全常识科普的目的在于提高人们的安全意识, 培养人们在意外来临时的自我保护能力与自救能力, 但是目前的安全常识科普方式不能有效反应各类安全常识之间的关系。论文依据安全科普网的数据, 基于Neo4j构建安全常识导图, 将数据以可视化的形式呈现, 建立以安全常识为核心的知识体系, 方便用户查询与学习安全常识。

关键词

Neo4j, 安全常识, 常识导图

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

安全问题是国际关注的重要问题之一, 是关系国家安全、人民健康、经济发展和社会稳定的重大战略问题[1]。安全教育是一种寓教于乐, 学有所获, 对培养创新型人才有辅助作用的教育方式, 具有教育人、培养人的社会功能[2]。安全知识科普和其他主题科普一样, 新颖科普形式容易达到良好效果。传统的安全常识科普基本是通过海报、新闻等媒介进行普及, 这种普及模式不能有效反应各类安全常识之间的关系, 不便于用户查询所需了解的安全常识。Neo4j 是一种图形数据库, 能快速对图形结构进行检索。论文基于 Neo4j 构建安全常识导图, 以数据可视化的形式展现各类安全常识之间的关系, 便于用户进行基于安全常识的检索。

2. 相关工作

安全问题引起国家的高度重视。2011年2月12日国务院学位委员会第28次会议通过《学位授予和人才培养学科目录》, 将安全学科列为一级学科, 从而确立了国家高等教育层面的安全人才培养的独立序列, 建立了安全人才长效培养机制[3]。2013年1月22日, 启动了《中国安全科学百科全书》编纂工作[3]。在安全科学迅猛发展的大背景下, 对研究进行系统总结和分析具有重要意义。当今经济发展迅速, 人们生活方式变得复杂多样, 危害生命安全的情况也随之增加, 因而安全知识科普成为日常学习的重要环节。但是对安全常识数据可视化研究较少, 安全常识数据量较大且关系复杂, 需要构建安全常识导图对数据进行系统化管理、可视化及检索。

Neo4j 图形数据库常用于数据系统化管理、可视化与检索。在文献[4]中, 廖理基于 Neo4j 实现了时空数据的存储。在文献[5]中, 陆鹏基于 Neo4j 实现了大数据的组织与检索。在文献[6]中, 张前进实现基于 Neo4j 智能学习系统语义链接图式存储研究。常识是在社会中生活所具备的基本知识, 包括生存技能、基本劳作技能、基础的自然科学以及人文社会科学知识等。常识导图是对常识节点和常识之间关系结构的可视化。目前, 对常识导图的研究较少, 主要研究集中在知识图谱方面。在文献[7]中, 朱木易洁提出知识图谱的构建主要包括自底向上(bottom-up)和自顶向下(top-down)两种方式, 目前大多数知识图谱都是采用自底向上的方式进行构建。知识图谱的构建主体包括知识获取、知识融合、知识加工和知识更新等

4个方面, 数据来源不同导致构建流程有所差异。

3. 安全常识导图构建

为了更好地反应安全常识之间的关系, 实现基于安全常识的检索。论文依据安全科普网(<http://www.aqkpw.com/>)的数据, 实现基于 Neo4j 安全常识导图的构建。

3.1. 安全常识获取

对安全科普网进行分析, 获取有用的安全常识数据节点进行分类, 构建安全常识导图的基本结构(如图 1 所示)。

安全常识包括职业健康、生产安全、心理健康、案例解密、科学救援、安全文化、民生安全和生命体验八类。民生安全包含了食品药品安全常识、医疗服务安全常识、教育安全常识、就业安全常识、环境保护常识、社会治安安全常识等重要的民生安全常识。职业健康包含了各职业常见的职业病及如何预防。生产安全介绍了在生产经营活动中, 为了避免造成人员伤害和财产损失的事故而采取相应的事故预防和控制措施, 使生产过程在符合规定的条件下进行, 以保证从业人员的人身安全与健康, 设备和设施免受损坏, 环境免遭破坏, 保证生产经营活动得以顺利进行的先关活动。生命体验通过情景模拟和现场体验把安全理论知识学习与实践操作技能有机融合的一种体验式安全教育模式, 侧重让受教育者在模拟的危险情境中接受安全知识教育, 提高应急反应能力, 提升安全防范意识和互救自救能力。心理健康主要介绍了减轻压力的措施和方法。案例解密对每次意外事故进行了详细报道。安全文化介绍了安全文化的作用以及如何加强安全文化教育。科学救援介绍了意外事件之后的救援工作。针对以上八类安全常识, 构建安全常识导图。

3.2. 常识导图构建与分析

论文基于 Neo4j 图形数据库构建了安全常识导图。Neo4j 是一个高性能的图形数据库, 通过节点(nodes)、关系(relationships)以及属性(properties)完整地描述任意一个用户的情况。在关系上赋予属性, 可以灵活扩展其网络模型;节点可以随时增删, 从而有效地解决存储半结构化、非结构化数据时内存浪费的问题, 通过其独有的深度(deep)遍历接口, 可以快速找到不同关联关系下的任意深度节点[8]。

常识导图构建的基本流程为信息抽取、常识融合和常识加工。信息抽取是从安全科普网中提取出节点、属性, 并获得节点之间关系; 常识融合是在获取新常识后对其进行整合以消除矛盾和歧义; 常识加工是对获取的新常识, 需要经过质量评估之后(部分需要人工参与甄别), 才能将合格的部分加入到安全常识库中, 以确保安全常识库的质量, 新增数据之后, 进行常识推理, 拓展现有常识, 得到新常识[9]。安全常识导图是由节点及关系、概念及其关系所构成的语义性常识库, 将安全科普的八类安全常识作为节点, 节点之间的关系作为边, 可以建立互联互通的常识导图(如图 2 所示)。

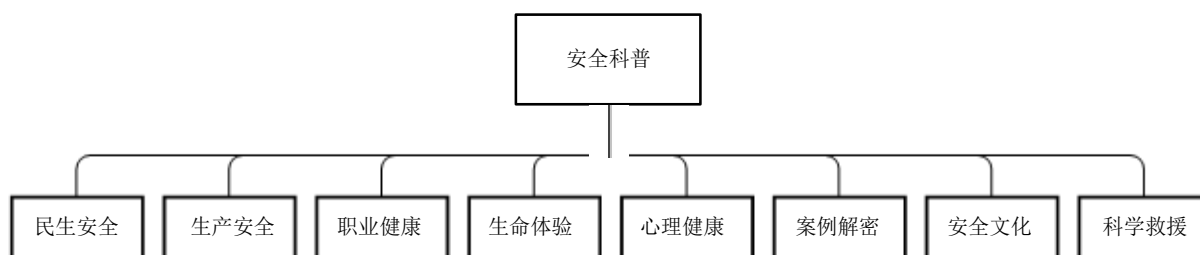


Figure 1. Safety common sense classification

图 1. 安全常识分类

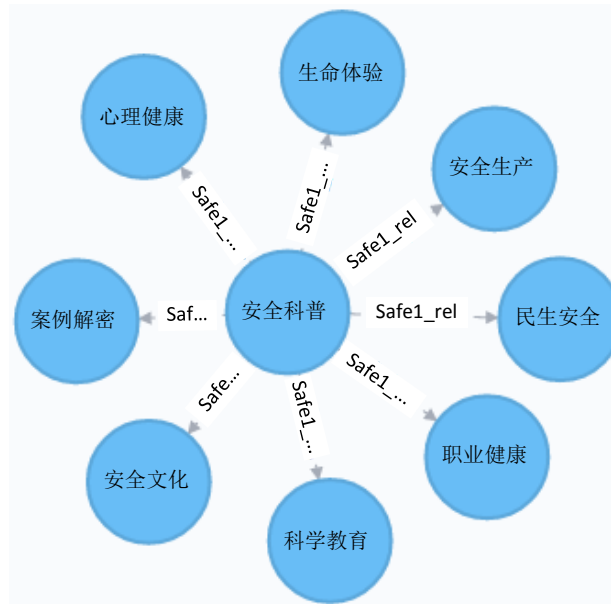


Figure 2. Security science common sense map
图 2. 安全科普常识导图

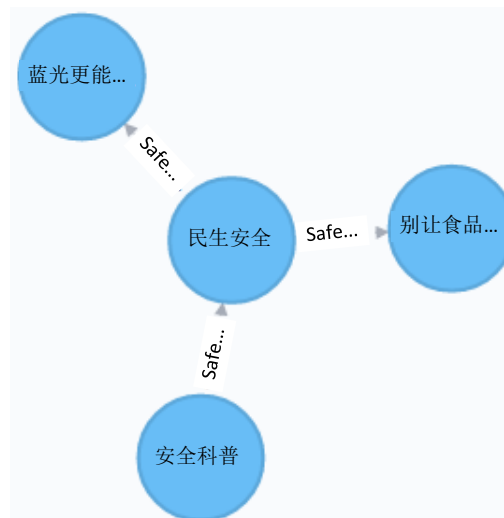


Figure 3. Livelihood safety common sense map search results
图 3. 民生安全常识导图查询结果

导图构建流程为:

Step1: 分别创建 node.csv 和 relationship.csv 两个文件, node.csv 代表所有的节点信息, relationship.csv 中的内容代表了节点之间的关系。

Step2: 基于 Neo4j 创建节点信息表 safe_test3。创建语句为
LOAD CSV WITH HEADERS FROM "file:///node.csv" AS line
MERGE (p:safe_test3{id:line.node_id,name:line.node_name})

Step3: 基于 Neo4j 创建关系表 safe_test3_rel。创建语句为
LOAD CSV WITH HEADERS FROM "file:///relationship.csv" AS line

```
match (from:safe_test3{id:line.from_id}),(to:safe_test3{id:line.to_id})
merge (from)-[r:safe_test3_rel{pro1:line.pro}]->(to)
```

根据构建的安全常识导图,既可以挖掘出潜在的关联关系,又可以对安全常识导图进行扩充。Cypher 是 Neo4j 提供的一种声明式、类 SQL、灵活且表达力强的查询语言,可以实现任意图谱的查询过滤[4]。

例如,查询民生安全常识,Cypher 语句为:

```
match (p:safe_test3 {name:"民生安全"})
return p。
```

查询结果如图 3 所示。

4. 总结

论文构建基本安全常识库,建立基于 Neo4j 的图形数据库模型,从而构建安全常识导图,对安全科普网的子页面及其相关内容进行常识提取。通过对数据可视化,实现了交互式查询服务,将各类安全常识关系进行梳理,更加易于人们学习和检索。

基金项目

国家科技支撑课题:安全科普和资讯聚合与推送系统及研发(课题号:2014BAH13F02)。

参考文献 (References)

- [1] 王宁,盛立. 生物安全相关文献知识图谱分析[J]. 军事医学科学院院刊, 2010, 34(4): 370-372.
- [2] 李力. 创新型开展“安全类”科普教育活动[J]. 青年时代, 2017(10).
- [3] 李杰,郭晓宏,吕鹏辉. 安全科学知识图谱的初步研究——以《SafetyScience》期刊数据为例[J]. 中国安全科学学报, 2013, 23(4): 152.
- [4] 廖理. 基于 Neo4j 图数据库的时空数据存储[J]. 信息安全与技术, 2015, 6(8): 43-44+56.
- [5] 陆鹏. 基于 Neo4j 的大数据组织检索研究与应用[D]: [硕士学位论文]. 南京: 东南大学, 2015.
- [6] 张前进. 基于 Neo4j 的智能学习系统语义链接图式存储研究[J]. 佳木斯大学学报(自然科学版), 2017, 35(2): 299-301.
- [7] 朱木易洁,鲍秉坤,徐常胜. 知识图谱发展与构建的研究进展[J]. 南京信息工程大学学报(自然科学版), 2017, 9(6): 575-582.
- [8] 康杰华,罗章璇. 基于图形数据库 Neo4j 的 RDF 数据存储研究[J]. 信息技术, 2015(6): 115-117.
- [9] 刘峤,李杨,段宏,刘瑶,秦志光. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-6753, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: jisp@hanspub.org