



开放科学
(资源服务)
标识码
(OSID)

煤矿标准文件知识图谱构建与应用

刘鹏^{1,2} 曹新晨³ 耿念³ 魏微³ 孟磊^{1,2}

1. 徐州高新区安全应急装备产业技术研究院 徐州 210000;
2. 矿山互联网应用技术国家地方联合工程实验室 徐州 221008;
3. 中国矿业大学信息与控制工程学院 徐州 221116

摘要: [目的/意义] 为了解决煤矿标准文件知识内容繁多、专业跨度大、组织模式有限,以及缺少系统、高效的煤矿标准文件知识存储利用手段,煤矿标准知识利用率低下等问题,以煤矿相关国家标准和行业标准文件为数据来源,构建了煤矿标准文件知识图谱。[方法/过程] 首先通过分析标准文件特点结构,提出了面向煤矿标准文件的标准文件-标准核心要素-标准条文的三层知识模型,定义了标准文件层、标准核心要素层以及标准条文层的知识单元结构和实体关系模式;然后根据设计的图谱模式层,针对半结构化和非结构化数据特点采用基于模板规则和基于深度学习的方法进行知识抽取,并选取图数据库存储方案,实现煤矿标准文件知识存储更新,完成了煤矿标准文件知识图谱构建。[结果/结论] 煤矿标准文件知识图谱可实现多种标准可视化方式以增强用户阅读体验,支撑智能检索问答应用帮助从业人员快速检索定位所需标准知识,还能提供标准冲突检测、辅助编写、性能指标比对等多种标准信息服务应用,显著提高了煤矿标准知识利用率,助力煤矿智能化建设。

关键词: 煤矿标准文件;知识图谱;知识抽取;标准信息服务

中图分类号: G35; TD67

Construction and Application of Knowledge Graph of Coalmine Standard Documents

LIU Peng^{1,2} CAO Xincheng³ GENG Nian³ WEI Wei³ MENG Lei^{1,2}

1. Xuzhou High Tech Zone Safety Emergency Equipment Industrial Technology Research Institute, Xuzhou 221000, China;
2. The National and Local Joint Engineering Laboratory of Internet Application Technology on Mine, Xuzhou 221008, China;
3. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China

Abstract: [Objective/Significance] Nowadays, many issues exist in application of coalmine standards, including complex knowledge, wide range of specialties, limited organizational patterns, low utilization efficiency and lack of systematic and efficient means for storing and utilizing coalmine standard documents knowledge. To address these issues, a knowledge graph

基金项目 江苏省安全应急装备技术创新中心(BM2022013);国家自然科学基金资助项目(71972176)。

作者简介 刘鹏(1973-),博士,副教授,研究方向为自然语言处理和知识图谱等;E-mail: liupeng@cumt.edu.cn;曹新晨(1996-),硕士研究生,研究方向为知识图谱构建;耿念(1996-),硕士研究生,研究方向为知识图谱构建;魏微(2000-),硕士研究生,研究方向为信息抽取;孟磊(1981-),博士,副研究员,研究方向煤矿知识图谱及煤矿智能化。

引用格式 刘鹏,曹新晨,耿念,等.煤矿标准文件知识图谱构建与应用[J].情报工程,2023,9(5):73-83.

of coalmine standard documents is constructed based on relevant national and industrial standards. [Methods/Processes] Firstly, by analyzing the characteristics and structure of standard documents, a three-layer knowledge model of D-E-P (Documents-Elements-Provisions) for coalmine standards is proposed. The knowledge unit structures and entity relationship patterns of the three layers are defined. Then, according to the schema of the knowledge graph, the knowledge extraction methods based on template rules and deep learning are adopted for semi-structured and non-structured data features, and the graph database was selected to realize the storage and update of coalmine standard knowledge. [Results/Conclusions] The knowledge graph of coalmine standard documents can be used to visualize multiple standards to enhance reading experience, support intelligent retrieval and Q&A applications to help practitioners quickly retrieve and locate required standard knowledge. It can also provide various standard information service applications such as standard conflict detection, auxiliary writing, and performance indicator comparison, significantly improving the utilization efficiency of coalmine standard knowledge and promoting the intelligent construction of coalmines.

Keywords: Coalmine Standard Documents; Knowledge Graph; Knowledge Extraction; Standard Information Service

引言

随着煤矿安全生产技术进步和国家政策推进,煤矿安全形势有所好转,但煤矿灾害仍处于高发态势,煤矿建设生产面临的安全隐患不容小觑^[1]。煤矿标准文件作为煤矿生产管理者必须遵循的最低要求,是煤矿能够安全建设的保证,更是政府部门监管的依据。目前,煤矿标准文件的数据信息服务仍处于初级阶段,多数煤矿企业仍使用传统的纸质版、电子版标准文件,标准文件数据库以 PDF、图片等形式存储,缺乏系统、高效的煤矿标准文件知识存储利用手段,不利于煤矿标准化、信息化发展。知识图谱是一种结构化的语义网络^[2],用于描述真实世界的概念、实体及其相互关系^[3]。知识图谱对知识单元及其间关系有良好的表达,提供了更好地组织管理海量信息的能力^[4],在语义搜索、问答系统、智能推荐等应用广泛^[5]。在煤矿领域,吴雪峰等^[6]构建了煤矿巷道支护领域知识图谱,有利于实现巷道支护智能化管理;曹现刚等^[7]构建了煤矿装备维护知识图谱,助力煤矿装备智能化动态管理;潘理虎等^[8]结

合本体技术构建了煤矿领域知识图谱,并开发了煤矿安全监测监控系统;李哲等^[9]构建了煤矿机电设备事故知识图谱,可应用于煤矿机电设备事故诊断、风险管理。

现阶段研究人员已将知识图谱应用于煤矿领域,但由于煤矿标准文件的数量内容繁多、专业跨度大,对煤矿标准的智能化研究少有涉足。而在通用标准文件领域,张慧等^[10]从标准文件的“前言”和“规范性引用文件”等共性要素入手,构建标准文献知识图谱。刘慧琳等^[11]使用知识图谱的模式对标准文件进行语义组织,提供了一种标准文件的知识图谱组织模式。郝文建等^[12]定义了标准文件基本要素和关键要素,分析了标准文件知识图谱构建方法。秦丽等^[13]分析了食品安全国家标准中三元组类型,构建了食品安全国家标准图谱。赵伟等^[14]解析了标准文件特点及结构,分析了标准文件的知识关联关系,构建了标准文件知识图谱 RDM 模型。李臻等^[15]提出了一种标准文件知识图谱构建流程,定义了标准要素实体和标准文件间关系,分析了标准文件知识图谱的应用场景。杨跃翔等^[16]分析了标准文件的共性结构要素的概念和关系,构

建本体层，将标准文件转换为 XML 文档解析，构建标准文件知识图谱。总体上，标准文件知识图谱构建的相关研究主要针对标准文件数据中共性要素的解析，对于标准文件的核心技术内容主要采用按章节结构分解的方式，缺少对标准条文内容的深度挖掘。

本文在分析煤矿标准文件特点、内容和结构的基础上，创新性地建立了“标准文件 - 标准核心要素 - 标准条文”三层煤矿标准文件知识模型；然后根据每层数据的特点设计不同的知识抽取方案，利用图数据库存储煤矿标准知识，完成了煤矿标准文件图谱构建；最后探讨了煤矿标准文件知识图谱的应用价值。

1 煤矿标准文件

1.1 煤矿标准文件概念及特点

标准文件的定义是通过标准化活动，按照规定的程序经协商一致制定，为各种活动或其结果提供规则、指南或特性，供使用者共同使用或重复使用的文件。广义的标准文件是标准化活动相关的所有文献资料。本文中煤矿标准文件指与煤矿生产、建设、安全等相关的国家标准、行业标准及其分类资料、检索工具等。

标准文件的内容、形式及适用范围等方面具有与一般文本不同的特点，具体表现在：

- (1) 编排规范。标准文件具有特定的编排格式和起草规则，文本整齐规范。
- (2) 协调性。一个标准化对象尽可能集中在一个文件中，标准文件间避免重复和不必要的差异。
- (3) 时效性。标准文件是标准化对象的底

线和门槛，随着经济的发展、科技水平的提高，标准文件会不断地修订、补充、替代或废止。

(4) 准确性。标准文件是规范性文件，其内容需专业、确切、规范，不应有错误和缺失。

标准文件前两个特点使其能够利用知识图谱技术进行结构化重组，同时也由于标准文件图谱的时效性与准确性，图谱的构建方法必须易于更新存储，同时能满足标准文件图谱的准确性需求。

煤矿标准文件主要包括煤矿国家标准（GB）、煤炭行业标准（MT）、能源行业标准（NB）和煤矿安全标准（AQ），标准类别包括基础标准、产品标准、试验标准、方法标准、管理标准等类型，涉及煤炭资源、煤矿地质、煤矿生产、煤矿运输贮存、煤矿设备、煤矿安全等专业内容，种类繁多，内容复杂，难以归纳定义完整的有普适性的煤矿领域本体模型。因此，煤矿标准文件知识图谱构建需要从标准文件的共性要素入手，构建标准知识图谱本体框架，技术要素等差异化内容依附于本体框架各自建立概念。

1.2 煤矿标准文件结构分析

煤矿标准文件由标准要素组成，根据 GB/T 1.1—2020《标准化工作导则第 1 部分：标准化文件的结构和起草规则》，依据要素的作用和位置，可将标准文件中的要素划分为规范性要素和资料性要素：规范性要素可划分为规范性一般要素和规范性技术要素；资料性要素可划分为资料性概述要素和资料性补充要素。标准要素的表述形式有条文、图、表、数学公式、示例、脚注、引用等。标准要素信息如表 1 所示。

表1 标准要素信息

要素类别	要素	存在状态	表述形式
规范性要素	规范性一般要素	范围	必备
		规范性引用文件	必备 / 可选
		术语和定义	必备 / 可选
		符号和缩略语	可选
	规范性技术要素	分类和编码	可选
		核心技术要素	必备
资料性要素		其他技术要素	可选
	资料性概述要素	封面	必备
		目次	可选
		前言	必备
		引言	可选
	资料性补充要素	参考文献	可选
		索引	可选

资料性要素是标准文件理解和使用的附加信息,其内容与格式相对统一、固定;规范性要素是界定文件范围或设定条款的要素,内容和格式相对繁多、复杂。以往的标准文件知识图谱的构建重点多在于对资料性要素和规范性一般要素的挖掘,而构建完整的煤矿标准文件知识图谱,数据内容须覆盖所有规范性要素。

标准层次依据文件内容的从属关系将标准内容划分为相对独立的单元,采用部分、章、条、段、列项和附录的形式。部分是针对一个标准对象的文件因需求或编制目的不同而编制成若干部分;章、条、段和列项是文件层次划分的单元,每一章均应有章标题,有下属层次的条应有条标题,层级之间逐级细分、层层嵌套。将标准层次标题作为煤矿标准规范性技术要素的层次实体建立层次概念,可保留标准的行文逻辑和标准条文的上下级关系。

标准条文是由条或段表述文件要素内容所使用的文字和文字符号,即标准正文内容,可划分为描述性条文和规定性条文。描述性条文

是对标准化对象的描述,例如定义、分类、型号、方法等;规定性条文是标准化对象的具体规定,例如要求、规程、指标等。

2 煤矿标准文件知识建模

煤矿标准文件知识模型的构建关键是知识单元的划分。通用文档知识图谱^[17]是对文本关键词及其关系的挖掘,只能反映文本的主题或提纲,造成的知识损失较大,而标准文件知识图谱的知识单元划分应尽量保证文件语义完整,知识单元的粒度应满足知识图谱检索与应用的需求。标准文件知识图谱中加入要素、条文等段落、句子的实例形式,同时对条文进行深度挖掘,表现全面完整的标准文件知识内容。根据标准文件的要素作用和编排层次的分析结果,本文创新性地解析煤矿标准文件知识图谱架构,将其划分为标准文件 SD (Standard Documents)、标准核心要素 SE (Core Elements of Standard)、标准条文 SP (Standard Provi-

sions) 三个层次的知识单元, 建立标准文件知识图谱 D-E-P (Documents-Elements-Provisions, 标准文件-标准核心要素-标准条文) 知识模型。

标准文件层与标准核心要素层、标准核心要素层与标准条文层之间为包含关系。煤矿标准文件知识图谱结构如图 1 所示。

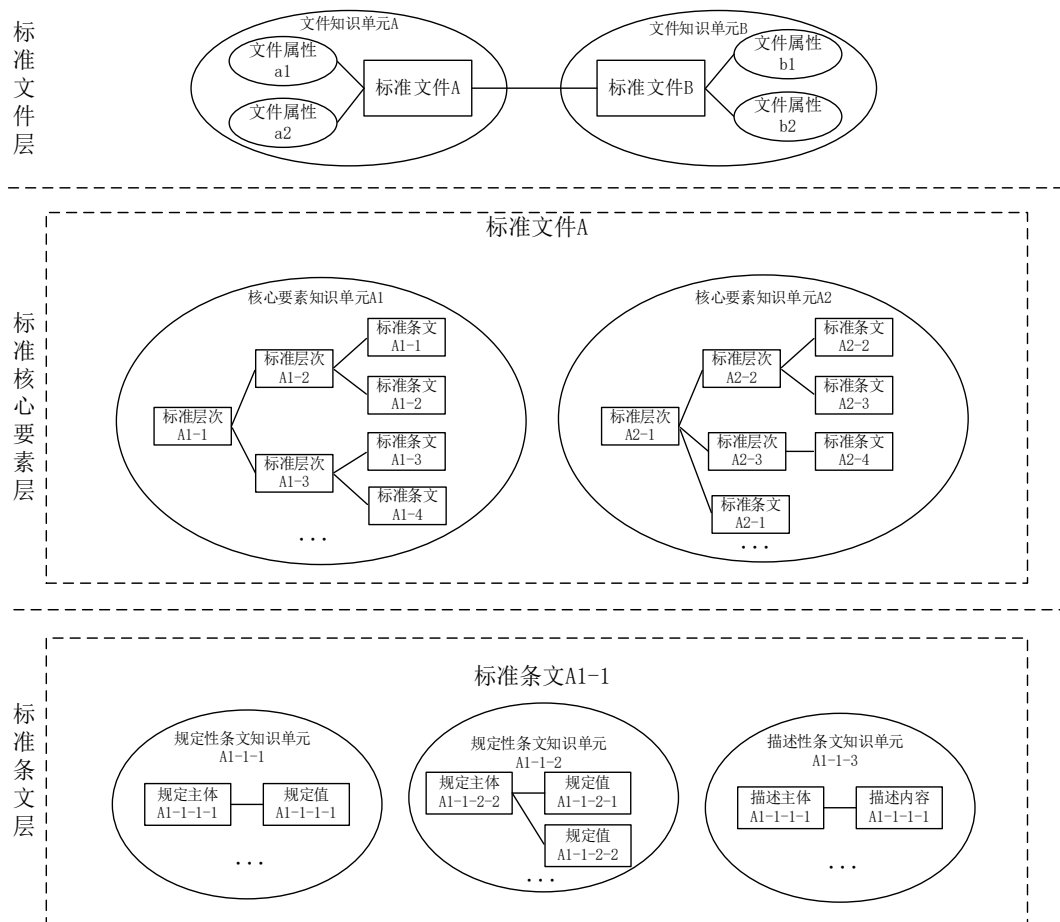


图 1 煤矿标准文件知识图谱结构

(1) 标准文件层

标准文件层以一个标准文件作为一个知识单元, 包含标准文件的全部属性和标准文件间的关系。标准文件层知识单元定义为 $SD=(D,C,R,A)$, 其中 D (Documents) 为标准文件实体; C (Concepts) 为标准文件相关实体对象的集合, 如人员、单位机构等; R (Relationships) 为标准文件间关系, 包括引用关系、替代关系等; A (Attributes) 为标准

文件自身属性, 如标准分类、发布日期、归口信息、适用界限等。

标准的资料性要素和规范性一般要素中包含大量标准文件属性和标准文件间关系, 标准文件层中主要概念、属性(关系)和所属要素如表 2 所示。以 GB 25974.1-2010《煤矿用液压支架 第 1 部分: 通用技术条件》(以下简称 GB 25974.1-2010) 为例, 标准文件层实例如图 2 所示。

（2）标准核心要素层

标准核心要素层以一篇标准文件下的一个核心技术要素作为一个知识单元，定义为 $SE=(D,L,P,R)$ ，其中 D (Documents) 为标准文件实体， L (Levels) 为标准层次实体， P (Provisions) 为标准条文实体，按照标准的“章”层次进行划分，“章”层次下的“条”或“段”作为标准条文实体，“章标题”和“条标题”作为标准层次实体； R (Relationships) 为实体间层次关系，包括上下位关系、继承关系等。标准核心要素知识单元中包含全部的标准层次实体、标准条文实体以及实体间的层次关系，保留了完整的核心要素知识和标准文件著者的原始逻辑。标准文件层次和实体如表 3 所示。以 GB 25974.1-2010 为例，标准核心要素层实例如图 3 所示。

表 2 标准文件层主要概念和属性(关系)

概念	属性（关系）	所属要素
标准文件名	标准编号	封面
	英文名称	
	中标分类	
	ICS 分类	
	替代文件	
	发布日期	
	实施日期	
	发布部门	
	起草依据	前言
	有关专利	
	归口信息	
	起草单位	
	起草人	范围
	标准化对象	
	适用界限	
	引用文件	规范性引用文件

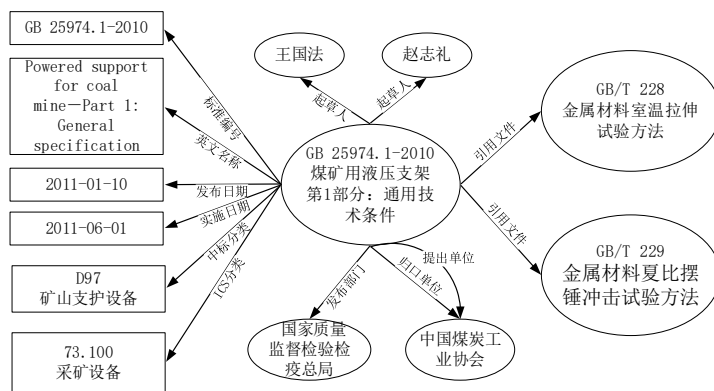


图 2 标准文件层实例

（3）标准条文层

标准条文层以一个标准条文句作为一个知识单元，对条文的标准化对象与标准化内容进行深度挖掘。标准条文层知识单元定义为 $SP=(P,V,S,R,A)$ ，其中 P (Provisions) 为标准条文实体，按照标准条文内容可分为描述型条文实体和规定型条文实体； S (Subject) 为标准条文主体，即该标准条文的标准化对象； V (Value) 为标准条文值，即该标准条文的标准

化内容； R (Relationships) 为标准条文关系，即标准条文主体与标准条文值间的关系，如表 4 所示； A (Attributes) 为标准条文属性，包括条文类型、所属要素层次等。

其中描述型条文包含描述主体和描述值，规定型条文包含规定主体和规定值。如 GB 25974.1—2010 中，条文“支架 powered support 以液压为动力实现升降、前移等运动，进行顶板支护的设备”为描述型条文，包含的三元组为：

(支架) - (英文名) -> (powered support) 、
(支架) - (定义) -> (以液压为动力实现升降、
前移等运动, 进行顶板支护的设备) ; 条文 “支
架内的行人通道的宽度应不小于 0.6m, 高度应
不小于 0.4m” 为规定型条文, 包含的三元组为:
(支架内的行人通道的宽度) - (范围) -> (不
小于 0.6m) 、 (支架内的行人通道的高度) - (范
围) -> (不小于 0.4m) 。标准条文层实例如图

4 所示。

表 3 标准文件层次和实体

层次	编号类型	实体
部分	第一部分	部分
章	1	章标题
条	1.1	条标题 / 条文
段	[无编号]	条文
条	1.1.1	条标题 / 条文
条	1.1.1.1	条文
列项	符号: “——” 和 “.” 编号: a) 和 1)	条文

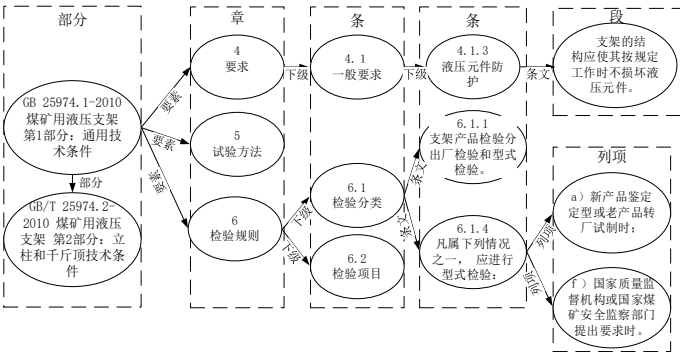


图 3 标准核心要素层实例

表 4 标准条文层关系模式

条文类型	标准化对象	标准化内容	关系	注释
描述型	描述主体	描述值	定义	标准化对象的定义
			分类	标准化对象的分类
			型号	标准化对象的型号
规定型	规定主体	规定值	符合	标准化对象与存在的式样、规定一致
			采用	标准化对象采用的方法、工具等
			具有	标准化对象具有的属性特征
			范围	标准化内容的界限
			防止	标准化对象阻止的事件
			应当	原则性规定内容
			必须	强制性规定内容
			可以	推荐性规定内容
			不得	否定性规定内容

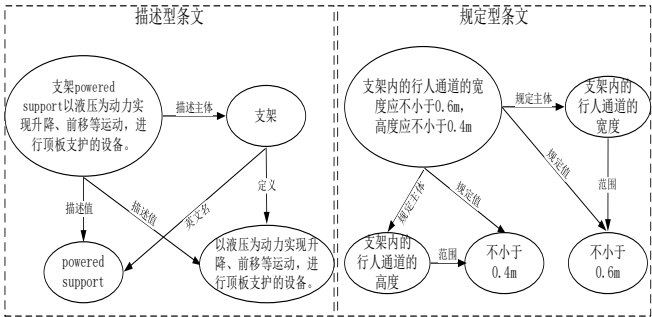


图 4 标准条文层实例

3 煤矿标准文件知识图谱构建

3.1 知识抽取

知识抽取是指从多种数据源中提取知识并存入知识图谱,是知识图谱构建的基础^[18]。煤矿标准文件的数据源包括半结构化数据、非结构化数据,针对标准文件模式层数据的特点采用不同的抽取方式可提高抽取效率,煤矿标准文件知识抽取流程如图5所示。

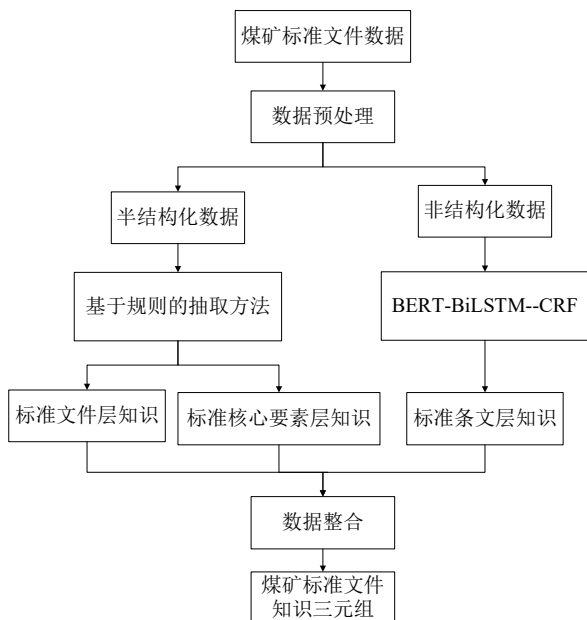


图5 煤矿标准文件知识抽取流程

(1) 数据预处理。在国家标准文件公开、行业标准信息服务等网站获取煤矿相关标准文件数据,并根据知识单元定义对数据进行划分。

(2) 半结构化数据抽取。标准文件层数据的结构规范书写逻辑严谨,标准核心要素层数据具有严格的层次编号,均可视为半结构化数据。在分析文件层各属性关系和标准核心要素结构位置信息的基础上,采用基于规则模板的抽取方法,构建标准文件层抽取规则,获取标准文件层知识及标准核心要素层知识。

(3) 非结构化数据抽取。标准条文层数据多为文本数据,没有固定的特征,是非结构化数据,构建规则模板成本高且泛用性低。因此,采用基于深度学习的方法进行抽取:将标准条文数据输入到 BERT-BiLSTM-CRF 模型中,利用 Bert 模型构成的字嵌入层将字转为词向量,利用 BiLSTM (Bi-directional Long Short-Term Memory, 双向长短期记忆网络) 层提取序列特征,利用 CRF (Conditional Random Field, 条件随机场) 层预测标签,得到标准条文层知识。

(4) 数据整合。对各层知识单元数据抽取结果进行整理,整合为 CSV 表格形式。为保证抽取数据的准确性,抽取完成后的结果需人工对抽取结果进行审查校正,以满足标准文件的准确性需求。

3.2 知识存储

知识图谱存储主要采用关系型数据库、RDF (Resource Description Framework, 资源描述框架) 和图数据库方式实现^[19]。关系型数据库需要大量表连接实现实体关系表示,运算开销大; RDF 三元组需事先制定特定领域 RDF 词汇表,且需借助特定 RDF 工具实现管理。而图数据库使用节点和边表示实体关系,且利用内置管理工具实现灵活、高效的存储和查询。因此,煤矿标准文件知识图谱选用 Neo4j 图数据库存储方案,其中标签对应标准文件图谱知识模型分层,节点及其属性对应实体及其属性,边对应实体间关系,存储方案如表5所示。

图数据库操作可利用 Neo4j 查询语言 Cypher 实现。本文使用 Python 中 py2neo 库执行 Cypher 语句实现数据库增删改查操作,常用

语句包括：Node、Relationship 类创建节点及关系；find、match 语句查询节点或关系；push 语句更新节点；run 语句执行 Cypher 语句等。

将获取的标准文件各层知识单元数据存储到 Neo4j 数据库后，形成的煤矿标准文件知识图谱共包含节点 16257 个，关系 22581 条。图谱数据统计信息如表 6 所示。

表 5 煤矿标准文件知识图谱存储方案

Neo4j 元素	作用	标准文件知识图谱对象
标签	根据标准文件知识模型对数据分层	标准文件层、标准核心要素层、标准条文层
节点	描述实体	标准文件实体、标准层次实体、标准条文实体、标准化对象实体等
属性	描述实体属性	标准文件属性、标准条文属性、标准化对象属性等
边	描述实体间关系	标准文件间关系、标准层次关系等

表 6 煤矿标准文件知识图谱统计信息

数据类型	节点个数	关系个数	关系类型
标准文件层	1023	1880	24
标准核心要素层	6972	16029	6
标准条文层	8262	4672	17
总计	16257	22581	47

表 7 知识图谱对比

图谱	知识来源	知识粒度	特点
煤矿领域知识图谱	煤矿领域开放知识	煤矿关键词为核心的较细粒度知识	针对煤矿特定领域应用，但缺少高质量数据来源，构建难度大。
其他标准文件知识图谱	标准文件	共性要素为核心的较粗粒度知识	高质量数据来源，但知识粒度较粗，构建方法效率低。
煤矿标准文件知识图谱	煤矿标准文件	标准条文为核心的细粒度知识	嵌套知识模型与细粒度数据内容，高效的构建方法，可满足多种图谱应用需求。

4.2 图谱应用

基于标准文件知识模型构建的煤矿标准文件知识图谱可以为标准制定、检索、阅读等提

3.3 图谱构建效果评估

我们已组织专业人士，针对煤矿标准知识图谱数据进行了均匀随机采样检测，对图谱质量进行了总体评估，并进行了针对性的修改优化，基本保证了图谱的专业性和准确性。需要指出的是，从业界目前研究现状看，图谱质量评估方法和指标体系尚处于研究初期，缺乏行之有效的有效的手段，实践可行性比较低。而且煤矿标准图谱构建是个复杂的系统工程，目前整体系统还在完善中，后续随着研究工作进展，我们将不定期邀请领域专家，从专业性、严谨性、易用性等角度对图谱进行持续质量评估。

4 煤矿标准文件图谱特点与应用

4.1 图谱特点

本文研究的煤矿标准文件知识图谱，以煤矿标准文件为切入点进行知识建模，创新性地提出并构建了“标准文件 - 标准核心要素 - 标准条文”三层知识单元模型，与其他知识图谱对比特点如表 7 所示。

供标准文件、标准核心要素、标准条文三种粒度知识单元的知识服务，有利于标准的推广应用。利用知识图谱相关技术开发的煤矿标准文

件知识图谱应用平台可实现标准可视化、智能检索问答、标准信息服务等应用。

(1) 标准可视化

传统的标准文件数据库以 PDF 图片或标准全文方式存储标准, 煤矿从业者在阅读多篇标准文件时难以找到所需的核心内容, 阅读过程中难以实现文档间的自由切换。标准文件知识图谱可视化可利用多种可视形式为用户提供更

好的阅读体验。通过标准文件知识图谱展示可以直接体现各标准知识单元间的关系; 利用数据分析方法将知识图谱内容以柱状图、饼状图等图表的形式展现出来, 有助于标准数据统计研究; 利用智能图书馆的文档表现形式可以全面系统地展示标准文件知识, 使用超链接跳转查阅方式实现用户的自由探索。标准可视化效果如图 6 所示。

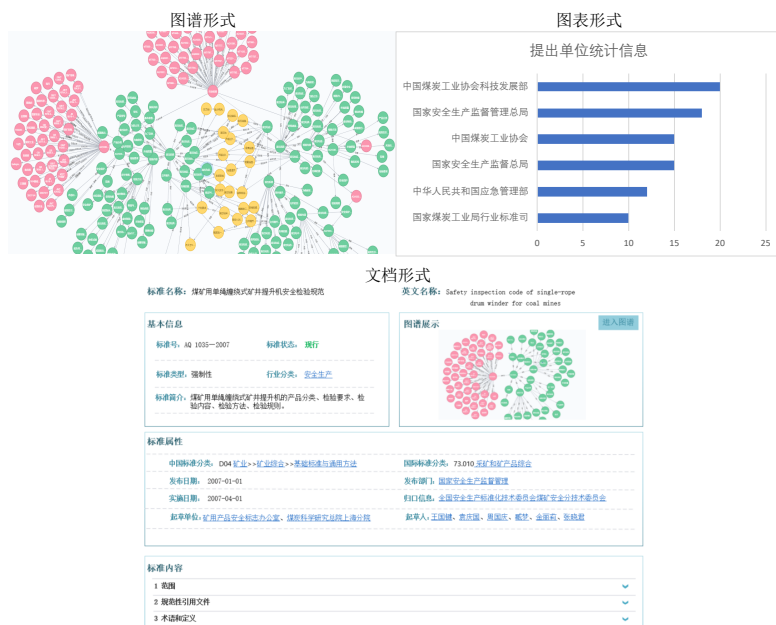


图 6 煤矿标准图谱可视化效果

(2) 智能检索问答

现阶段标准文件检索系统多数只能实现对标准名称、标准编号的简单检索, 搜索过程和结果中缺少标准文件内部信息, 致使煤矿从业者难以快速、准确地定位所需标准文件, 耗费时间精力研读标准全文寻找所需信息。借助标准文件知识图谱开发的智能检索问答系统可以为用户提供多级的、深层次的检索方式, 利用智能问答技术理解用户检索语义和意图, 提供自然语言问答交互功能。智能检索问答应用流程如图 7 所示。

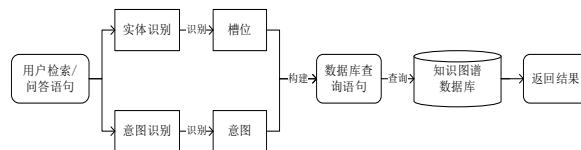


图 7 智能检索问答应用流程图

(3) 标准信息服务

利用标准文件知识图谱重组标准文件后, 可为煤矿标准化活动提供多种标准信息服务。条文内容的细粒度抽取可将多次出现的关键知识共享重用, 发掘标准文件的公共性和一致性, 同时能对现有标准内容进行冲突性检测; 利用高质量标准形成标准文件模板辅助标准文件编

写,为标准编写者提供涉及标准相关信息参考,保证标准的准确性;利用自然语言处理技术实现实际指标与标准文件图谱内容比对,辅助煤矿标准实施监督检查、达标检测、指标评价等活动,提高监督人员效率,指标比对流程如图8所示,指标比对效果如图9所示。

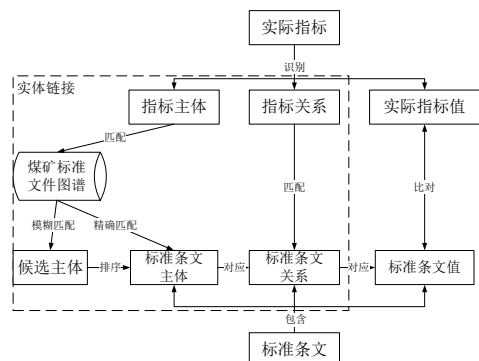


图8 指标比对流程



图9 指标比对效果截图

5 结语

本文在解析标准文件要素、层次和条文的特点的基础上,创新性地提出并构建了包含标准文件、标准核心要素和标准条文的三层知识单元模型,将标准文件知识进行重组,保留了标准要素内容的层次逻辑和复杂语义,对标准条文知识进行了细粒度挖掘;依据标准文件知识模型,通过规则模板和深度学习方法对不同类型数据进行了知识抽取,存储到 Neo4j 图数据库中,完成煤矿标准文件知识图谱构建;探讨了煤矿标准文件知识图谱的应用价值,为煤

矿标准化、数字化发展提供了可借鉴思路。后续笔者计划在煤矿标准文件知识图谱内容的融合更新及深度应用方向开展更进一步研究。

参考文献

- [1] 沙迪,李雨成,田叶,等.煤矿安全生产事故统计分析预测研究[J].高技术通讯,2018,28(1): 83-89.
- [2] 漆桂林,高恒,吴天星.知识图谱研究进展[J].情报工程,2017,3(1): 4-25.
- [3] AMIT Singhal. Introducing the Knowledge Graph: things, not strings[R]. US: The official google blog, 2012.
- [4] 刘峤,李杨,段宏,等.知识图谱构建技术综述[J].计算机研究与发展,2016,53(3): 582-600.
- [5] 王萌,王昊奋,李博涵,等.新一代知识图谱关键技术综述[J].计算机研究与发展,2022,59(9): 1947-1965.
- [6] 吴雪峰,赵志凯,王莉,等.煤矿巷道支护领域知识图谱构建[J].工矿自动化,2019,45(6): 42-46.
- [7] 曹现刚,张梦园,雷卓,等.煤矿装备维护知识图谱构建及应用[J].工矿自动化,2021,47(3): 41-45.
- [8] 潘理虎,张佳宇,张英俊,等.煤矿领域知识图谱构建[J].计算机应用与软件,2019,36(8): 47-54.
- [9] 李哲,周斌,李文慧,等.煤矿机电设备事故知识图谱构建及应用[J].工矿自动化,2022,48(1): 109-112.
- [10] 张慧,侯霞.基于知识图谱的标准文献分析[J].计算机工程与设计,2017,38(4): 1103-1109.
- [11] 刘慧琳,牛力.标准文件的知识图谱组织模式探究[J].档案学通讯,2021(5): 58-65.
- [12] 郝文建,魏梅,张浩,等.标准知识图谱的构建与应用[J].信息技术与标准化,2021(8): 44-47.
- [13] 秦丽,郝志刚,李国亮.国家食品安全标准图谱的构建及关联性分析[J].计算机应用,2021,41(4): 1005-1011.
- [14] 赵伟,张览,望俊成.标准文献知识图谱构建的模型设计与集成方法[J].情报工程,2021,7(6): 58-66.
- [15] 李臻,刘彦林,马小雯,等.基于标准知识图谱构建前瞻性研究[J].标准科学,2021,571(12): 63-68.
- [16] 杨跃翔,涂新雨,刘文玲.标准文献知识图谱构建与应用研究[J].数字图书馆论坛,2022,217(6): 22-30.
- [17] 杨建梁,刘越男,祁天娇.文档数据化:概念、框架与方法[J].中国图书馆学报,2022,48(3): 63-78.
- [18] 张吉祥,张祥森,武长旭,等.知识图谱构建技术综述[J].计算机工程,2022,48(3): 23-37.
- [19] 王鑫,邹磊,王朝坤,等.知识图谱数据管理研究综述[J].软件学报,2019,30(7): 2139-2174.