

基于绿色建筑设计策略与技术的知识图谱 搭建路径研究^{*}

官晓晴¹, 姜 涌¹, 宋晔皓¹, 张成英¹, 褚英男¹, 曹晓昕²

(1. 清华大学建筑学院, 北京 100084; 2. 中国建筑设计研究院有限公司, 北京 100032)

[摘要]建立垂直领域知识库、实现搜索优化与知识推动、自动知识生成与辅助决策,有助于推动传统建筑行业向数字化、智能化转型。梳理了建筑行业知识管理的需求和现状,结合知识图谱的发展历史,展望了建筑知识图谱的宏观图景。以绿色建筑设计策略与技术为例,提出一种面向特定需求的建筑知识图谱的建立路径,探讨了搭建过程中框架设计、数据标注和扩充的关键环节,并展望了建筑领域知识图谱的应用和发展前景。

[关键词]知识图谱;绿色建筑;设计;知识地图;人工智能;机器学习;语义识别

[中图分类号] TU204 **[文献标识码]** A **[文章编号]** 2097-0897(2025)22-0115-06

Research on Construction Path of a Knowledge Graph Based on Green Building Design Strategy and Technology

GUAN Xiaoqing¹, JIANG Yong¹, SONG Yehao¹, ZHANG Chengying¹,
CHU Yingnan¹, CAO Xiaoxin²

(1. School of Architecture, Tsinghua University, Beijing 100084, China; 2. China Architecture Design & Research Group, Beijing 100032, China)

Abstract: Establishing a vertical domain knowledge base, realizing search optimization and knowledge promotion, automatic knowledge generation and auxiliary decision-making, will help promote the digital and intelligent transformation of the traditional construction industry. Based on the necessity and current situation of knowledge management in the field of architecture, this paper looks forward to a macro picture and an application scenario of architectural knowledge graph. Taking green building design strategies and technologies as an example, it tries to put forward a path to build an architectural knowledge graph for specific needs. Key steps of this process, including top-level framework building, data input protocol and automatic data expansion are specifically discussed.

Keywords: knowledge graph; green building; design; knowledge map; artificial intelligence; machine learning; semantic recognition

0 引言

建筑物作为居住环境产品和大众消费品,与一般工业产品不同,具有耐用消费品和不动产资产属性。其工程设计、采购生产、运营维护涉及建筑物全生命周期的全过程、全参建方、全要素,构成了建

筑学庞大而复杂的知识体系。

建筑学是典型的知识生产型行业。建筑师在建筑生产中身兼数职,既是工程技术整合者、建设方顾问,也是项目管理者、行政许可申报者和施工监管者,其专业活动涉及的知识领域广泛,远超单个团队或企业的能力范围。建筑工程、知识体系及建筑师工作的复杂性,凸显了知识管理在建筑学中的重要地位。

1 作为知识管理核心工具的知识图谱

1.1 知识图谱的概念

数据、信息和知识是认知的3个层次。数据是

* 十四五国家重点研发计划:高品质绿色建筑设计方法与智慧协同平台(2022YFC3803800);住房和城乡建设部2022年度软科学研究项目:基于工序单元的建设工程质量管理的数字化平台关键技术研究和示范应用(R20220452)

[作者简介] 官晓晴,博士研究生, E-mail: gxq22@mails.tsinghua.edu.cn

[通信作者] 姜 涌,副教授, E-mail: Jyong2000@163.com

[收稿日期] 2025-03-10

原始、未经处理的数字或符号;信息是有意义、结构化且指向特定事实的数据;知识则是对数据和信息加工、结构化后,形成的认知、价值判断,能指导决策。

知识管理(knowledge management)旨在借助特定平台或工具,优化知识创造、分享、整合和更新的过程。20世纪70年代起,人工智能研究试图通过语义网(semantic web)、资源描述框架(RDF, resource description framework)、网络本体语言(OWL, web ontology language)等技术、框架和协议,将人类信息转为机器可理解内容。同时,随着信息存储技术和各类数据库工具的发展,面向大数据的知识图谱的建立成为了可能。

2012年,谷歌发布知识图谱项目,是知识图谱发展历程上的一个重要节点。它以“节点”对应概念和实体,“边”表示语义关系,抽取网页信息,关联多源异构知识库,构建统一结构化语义网络。在其支持下,搜索引擎实现从文本匹配到语义检索的转变,实现了检索效率大幅提升^[1]。

1.2 知识图谱搭建的基本路径

从最早的人工搭建,到基于百科数据库的自动搭建,再到近年来出现的基于开放域知识抽取的项目(如KnowItAll,NELL),知识图谱搭建技术随着信息技术的进步而不断优化,表现为效率提升、处理异构数据的能力逐步增强。虽然搭建技术多种多样,但知识图谱的搭建路径可以概括为2种基本模式:自上而下模式和自下而上模式^[2]。自上而下模式是先搭建顶层结构框架,然后将抽取到的实体与顶层结构进行匹配、归类。自下而上模式是利用多种抽取技术获得知识源中的实体、属性和关系,并将其置信度高的结构合并形成结构,后续可通过自然语言处理(NLP,natural language processing)等方法自动搜集、构建、扩充知识图谱(见图1)^[3]。

知识图谱按照专业程度可分为通用知识图谱和垂直知识图谱(或称行业知识图谱)。通用知识图谱强调知识的广度,通常使用自下而上的方式搭建,有利于自动生成知识并发现新的知识结构;垂直知识图谱可以看作基于语义技术的行业知识库,强调知识的深度和整体架构,通常采用自上而下和自下而上相结合的方式搭建。完全自上而下搭建的图谱则相对罕见。

2 建筑行业知识管理现状

2.1 工程信息管理

建筑工程全生命周期产生了大量的数据信息。为了建立良好的数字化协作环境,行业内目前已有

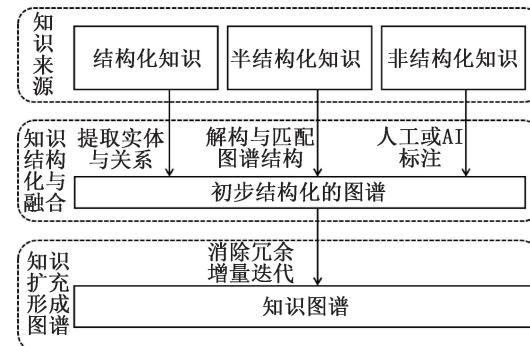


图1 自下而上搭建知识图谱的模式

Fig. 1 Bottom-up model for building knowledge graphs

多种信息交换和分类标准,例如北美的Omniclass分类体系,国际标准化组织的ISO 12006—2标准和我国的GB/T 51269—2017《建筑信息模型分类和编码标准》。这些信息分类和编码系统提高了工程管理效率,为建筑信息模型(BIM)提供了底层的数据管理框架。但数据的分类和标准化离建筑工程知识的结构化体系还有相当大的差距。除了工程数据之外,案例介绍、部品部件、图纸模型、法律法规、标准图集各类与建筑工程实践相关的信息分散于不同网站和知识库,彼此缺乏关联,导致行业知识传播与更新相对缓慢。

2.2 学术知识地图

知识地图(knowledge map, atlas of science)是知识可视化的范畴。在建筑学领域中,设计知识图示、学术动态分析、学术谱系研究等场景,都常用到知识地图工具。

1)设计知识图示常见于各类设计资料集,针对某个板块的设计知识进行初步结构化,并绘制相关图示,对建筑实践具有较强的指导意义。

2)学术动态知识地图,通过对特定领域学术论文关键词词频、突现词词频及出现时间、阅读量等指标进行统计和可视化,可以直观呈现该领域的研究脉络和主要研究内容,提示现状研究空白或未来热点方向,是学术研究动态分析的常用工具。例如陈翠等对“历史建筑商业化”这一主题的研究,通过学术动态分析,指出该领域“价值认识不足”和“理论指导失效”两大问题^[4]。

3)学术谱系知识地图,致力于描述某类或某几类实体(如建筑师、建筑学术著作、建筑作品)的学术网络关系。如江嘉玮对岭南学派传承谱系的研究,围绕不同阶段的“舶来与本土”特征进行分析与观察,为岭南学派内涵和外延的探讨提供新的思路(见图2)^[5]。

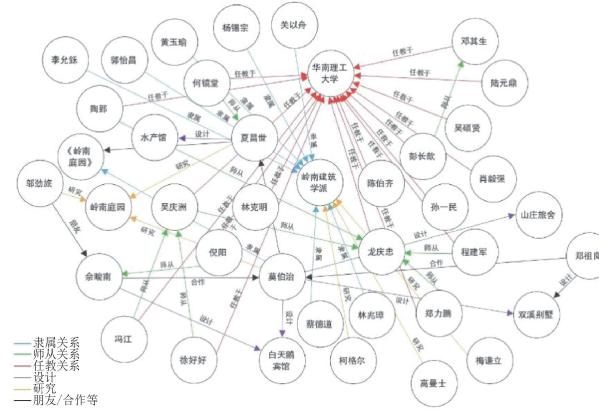


图 2 岭南学派社会网络关系的知识图谱研究

Fig. 2 A knowledge mapping study of social network relationships of Lingnan scholars

知识地图具有一定的结构化特征,可以理解为知识图谱的初级形态,有时也被视作广义的知识图谱。知识地图的强索引性,使它通常可以作为知识图谱的交互界面存在。但相比于狭义的知识图谱,知识地图的主观性较强、知识涵盖范围较小、结构化程度较低,拓展性不足,并不能指向行业知识体系建构与发展的智能化目标^[6]。

3 绿色建筑设计策略与技术图谱的搭建探索

3.1 建筑行业知识图谱的搭建路径

如上所述,建筑领域已有一定程度的知识管理,但目前尚未出现系统性的、可拓展的行业知识图谱。搭建建筑行业知识图谱是一项宏大的工程,不可能一蹴而就,可以从某个特定子领域的特定知识需求开始进行搭建的探索尝试。本文选取与当下“双碳”目标密切相关的绿色建筑领域,针对设计阶段建筑师辅助决策的特定需求,尝试通过“自上而下”与“自下而上”相结合的路径,建立绿色建筑设计策略与技术的知识图谱。具体步骤在杨玉基等^[7]提出的四步法基础上进行了合并与简化(见图3)。

1) 框架设计。根据专业知识体系建立顶层框架,即知识图谱的“本体”。本体构建可采用人工、自动和半自动的方法。顶层框架需要兼有系统性、宽容度和扩展性。

2) 数据标注。人工录入案例、技术、产品等信息完善领域本体,形成初步结构性的图谱。这一步需要建立合理的数据协议,解决建筑工程数据来源多样、数据结构差异较大的问题。

3) 数据扩充。选取可信任的开放数据源获取大量数据,按照领域知识图谱的本体结构,对这些数据进行结构化处理,补充领域知识图谱的实体和关系。

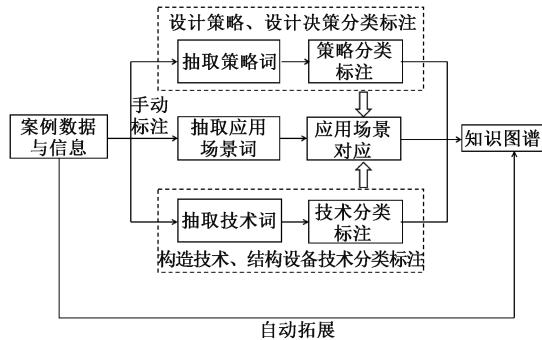


图 3 绿色建筑设计策略与技术图谱的搭建路径

Fig. 3 Path of building knowledge graph for green building design strategies and technologies

4) 信息抽取。采用有监督、半监督和无监督的方法从外源数据中抽取实体和关系扩充知识图谱,同时结合行业特点,可灵活采用分级搭建与嵌套融合的策略。

其中,前两步奠定了知识图谱的核心结构,是需要反复斟酌和反思的探索过程,后一步对知识图谱的扩充和丰富是建立在这一结构上的半自动技术过程。

3.2 框架设计——围绕场景、策略与技术的框架

绿色建筑设计知识体系庞杂,包括生产体系、评估体系、建筑本体系统,以及与政府相关的行业监督相关内容等(见图4)。从辅助建筑师前期决策的需求出发,首先应建立一个具有较强的系统性和较高的结构化程度的顶层框架,聚焦建筑本体系统,建立策略、技术以及应用场景的实体关系,后期可借助自然语言处理和机器学习等技术快速进行内容拓展,完善知识体系。

从辅助建筑师决策的角度出发,选择合适的策略和技术是建筑师及其工程团队的决策目标,而场景则包括了决策时所需要考虑的多种条件。因此,在搭建一个围绕绿色建筑设计知识的知识图谱时,以场景、策略、技术作为主要的节点类型,可以对现有散落的工程经验和知识进行快速的结构化(见表1)。场景既包括建筑的基本规模条件,也包括项目的功能类型、产权归属、投资性质、绿建目标等前置条件。策略则可以囊括建筑、暖通、水电等专业的各类决策。技术节点所包括的内容比策略更少,且更为具体,例如某一种构造做法,施工方法,乃至部品或材料的选择。将策略和技术进行区分,是因为二者处于决策流程的不同环节,具有不同的参考意义。知识图谱中的“边”则将不同的策略、技术与场景关联起来,形成一个面向决策辅助的知识图谱(见表2)。

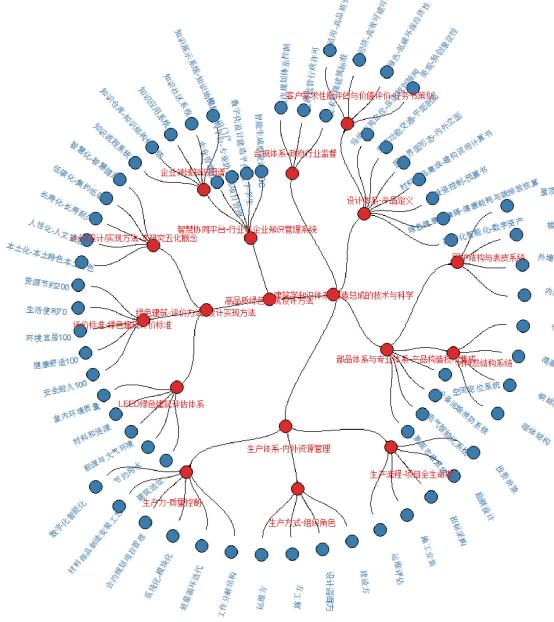


图 4 围绕绿色建筑设计方法的不同知识分类方式
Fig. 4 Different ways of categorizing knowledge around green building design

表 1 绿色建筑策略与技术图谱的实体类型与样例

Table 1 Entity types and samples of knowledge graph for green building design strategies and technologies

实体类型	实体子类型	样例
场景	建筑结构规模	多高层场景、大跨建筑场景
	建筑功能	商业、办公、教育、体育、交通、医疗
	建筑封闭程度	全封闭、半开放、开放
	是否改造	是、否
	绿色目标	低能耗、低碳排放、轻量化、长寿耐久、达到特定绿色建筑评级
策略	材料体系	钢结构体系、钢筋混凝土体系、木结构体系、混合结构体系
	结构体系	梁柱体系、桁架体系、墙承重体系、核心筒体系
	体量组合	向心型、放射型、线性延伸型、不规则型、整体型
	空间组织	整体式、单元组合式、包围层级式
	主被动设计策略	屋顶结构整合自然采光设计、控制用能空间体积以降低能耗
技术	设备集成策略	结构设备一体化设计
	构造技术	光伏组件墙体构造技术
	围护更新技术	钢木节点更换技术
	施工技术	现代木结构曲面屋面施工技术

3.3 数据标注——现代木结构案例信息库

现代木结构案例信息库是策略图谱搭建过程中的一个分支,其目的是建立一个可供建筑师检索参考的精选案例库,本质上是一个可以独立搭建、独立使用的知识图谱。同时,从中提取相应的技术和策略,也成为上述绿色建筑设计策略与技术知识图谱的核心内容。

表 2 绿色建筑策略与技术图谱的关系类型与样例

Table 2 Types and samples of relationships between entities in knowledge graph for green building design strategies and technologies

关系	示例
场景-场景	轻量化-旧建筑改造场景 大跨度-交通建筑场景
场景-策略	低碳排放-木结构体系 多高层-核心筒体系-混合结构体系
场景-技术	零能耗-光伏组件墙体构造技术 长寿耐久-钢木节点更换技术
策略-技术	混合结构体系-混凝土 CLT 复合楼板技术 空间网架体系-现代木结构曲面屋面施工技术
策略-策略	折板结构体系-自然采光策略 空间网架体系-照明设备一体化设计策略

案例的集合在建筑领域以多种多样的形式存在,因此现代木结构案例信息库的数据存在典型的多源异构的特征,即数据来源多样,不同来源的数据结构差异较大,不同来源数据的同义重合较多。搭建案例信息库的核心是要建立数据录入标准,保证结构化,消除冗余。从建筑案例相关的信息类型来看,现代木结构案例信息库的数据分为以下 4 类^[8-10]。

1) 第 1 类,建筑项目需要录入的常规信息,如建成时间、建设地点、各个分项的设计主体、建设规模等;此类信息在源数据中相对容易辨识,入库时需要对格式(时间格式、地点格式)和单位(面积单位)进行统一,并区别相近的概念(如建筑面积、用地面积)。需要特别注意的是,建筑面积、用地面积、容积率、绿地率、建筑密度等指标间存在相互换算的关系,可用于补全缺失的数据。

2) 第 2 类,根据策略与技术图谱的关注点收集对应的基础信息,如结构规模分类、结构体系、节点体系、围护结构做法、主要跨度数据、主要结构构件截面尺寸、木材类型、树种与规格信息等。此类信息需要进行标签化处理。部分标签来源于数据内部,且具有一定的行业共识,标注难度较低,例如结构体系、木材类型的相关标签。部分标签来自外部,例如“结构规模分类”,下属的 3 个标签“单元系统”“跨度系统”“多高层系统”,则具有一定的主观性和功能性,是在客观的结构跨度、建筑层数基础上,结合了建筑师关心的空间特征、设计重心、结构表现力等定义的标签,可以方便使用者进行案例条目的筛选。在大规模建筑行业知识图谱的搭建过程中,此类标签的定义需依赖专家系统的介入。

3) 第 3 类为相对难以结构化的文字描述信息,如案例介绍、使用者反馈、设计方访谈、论坛评价

等,其中包含了大量与设计决策、技术创新与策略组合相关的内容,其标签的定义和分类均存在一定的难度。人工标注时,采用“提取信息-合并相似选项-修改标签表述”的方式,通过一定量的数据录入,建立1个基础的标签系统和分类原则。后期录入时,则需要通过自然语言处理技术和有监督机器学习,实现非结构化的文字信息与既有标签的自动识别、自动对应。

4)第4类信息为图片附件。目前仅采用人工标注的方式进行了粗略分类,未来可借助人工智能技术自动进行图片类型的识别,与策略、技术类节点关联匹配。

第1,2类信息解决了数据结构差异的问题,第3,4类信息则具有一定的拓展性,保证了信息录入的完整性、丰富性和可读性,而标注难度也较前两类信息更高。

3.4 数据扩充——分级搭建与嵌套融合

上述搭建建筑学领域特定主体知识图谱的过程显示出,围绕场景、策略、技术3类实体建立的建筑行业知识图谱领域具有一定的合理性和应用价值,后续通过网络数据自动爬取、自然语言处理、机器学习标注的方式实现自动扩展也具有理论可能性。然而,人工标注过程已经提示了后续外源数据补充与自动标注过程中可能存在的问题。

一方面是领域本体构建阶段,顶层框架难以达成学科内部共识的问题。建筑学的知识体系中存在大量的隐性知识,许多与实体命名、标签分类的认知,往往不来源于文本内部,上述现代木结构案例信息库中的“结构规模分类”就是例子。隐性知识的存在决定了建筑知识图谱的领域本体难以通过自下而上的方式自动生成,而是需要领域专家的介入。对于大规模的行业知识图谱而言,就一个顶层结构达成行业共识,在操作上具有较大难度。

另一方面,建筑领域的术语和语汇体系具有较高的复杂性,一定程度上会对语义识别造成阻碍。有的术语含义丰富,需要依赖语境识别其意义,典型例子如“建构”这一术语,在学术语境中常用“tectonic”的含义,在工程语境中经常用于表达“build”(搭建)的意思;有的信息可能对应截然不同的多种文字表述方式,但实际归属同类,在合并标签时容易遗漏,或创建标签时容易产生冗余和重复。

综上所述,目前通过人工标注与自动标注相结合的方式,创建特定主体的小规模知识图谱并非难事。但在建立更大范围的行业知识图谱,尤其是将建筑本体之外的生产体系、政策法规体系、评估评

价体系的知识纳入知识图谱时,还是会面临相当大的操作难度。

因此,当前建筑学领域知识图谱的搭建,在“自上而下”与“自下而上”相结合之外,分级搭建、嵌套融合也是非常重要的思路,即由多个不同层级的小型图谱对接、拼合,最终形成相对完整的学科知识图景。上述绿色建筑设计策略与技术知识图谱提供了一个分级搭建、嵌套融合的例子。现代木结构案例信息的录入,是内容界定明确、核心工作量可控的过程,其录入标准在有限次的迭代调整之后便趋于清晰,可相对快速地进入数据扩充阶段。围绕场景、策略、技术的本体建构具有一定的复用性,可以应用到装配式钢结构建筑案例、现代竹结构案例等相似领域主题中。其手动信息录入过程提取了策略词和技术词标签,为绿色建筑策略与技术知识图谱提供了初始的实体内容,后续可基于这些标签进行快速的内容扩展。

知识图谱的结构虽然具有可视化的形态,但因为实体和边的数量众多,通常会造成阅读、索引上的困难。在知识图谱的基础上,围绕特定主体提取知识图谱的相关内容,形成强索引功能的知识地图,也是知识图谱的一大应用前景(见图5)。

4 结语

建筑行业作为传统的国民经济支柱产业,涉及的参建方和专业领域庞杂,每个技术策略的选择与优化都需要多维度、多视角的比较权衡。建筑行业目前的知识管理还处于结构化分类的初级阶段,决策过程通常由案例检索、供应商人脉、参考图集等方式碎片化地提供技术参考。引入专业化的知识管理,是学科知识积累、传播与迭代的基础,也是行业效率优化、生产模式转型、建筑品质提高的关键,具有重要意义。

本文以绿色建筑设计这一特定主题、快速辅助决策这一特定需求,借助一个小型图谱的搭建案例,探讨了顶层架构设计、底层数据标注与数据扩展3个关键步骤,初步探索了建筑行业知识图谱的搭建路径。目前,围绕特定主题,采用自上而下与自下而上相结合的方式进行小型知识图谱的搭建,已具备成熟的条件。难点在于,随着主题的扩大,顶层框架的设计具有一定的主观性,难以取得大范围的学术共识;随着生成式人工智能技术(AIGC)和语义识别技术的不断发展,为通过纯技术手段进行自动知识提取、顶层框架融合带来了可能性。

可以预见,建筑行业知识图谱在人工智能技术的辅助下具有广阔的应用前景:支撑专业检索引擎

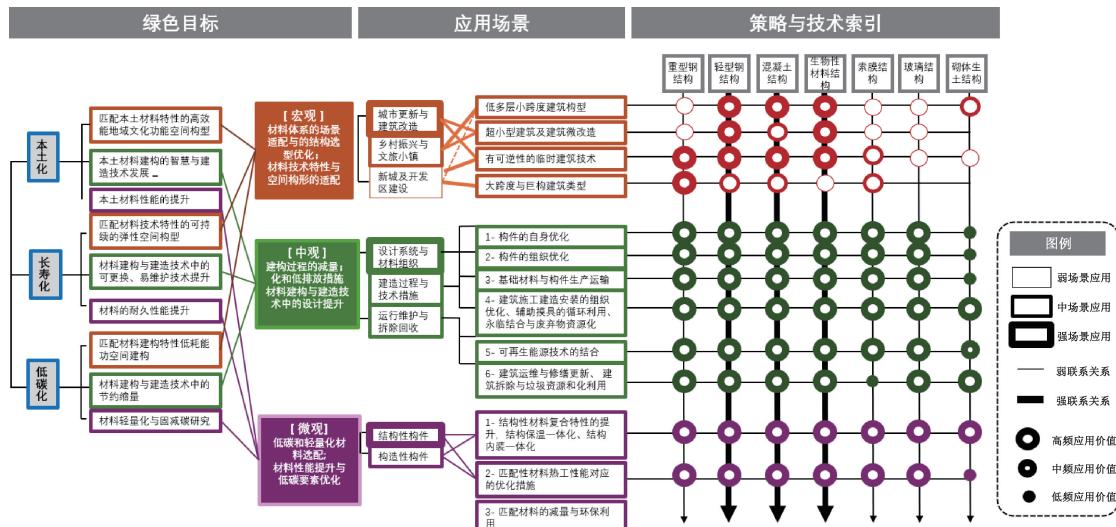


图 5 绿色建筑设计策略与技术索引知识地图示例

Fig. 5 Example of an index map of green building design strategies and techniques

的开发,根据建筑师的职业需求和不同的工程阶段,对检索结果进行优化排序或智能推送,提高知识转化效率;基于建筑行业知识图谱开发智能推荐算法,向目标用户推广与近期项目密切相关的行业知识、法规、案例、技术措施建议、产品信息等,并实时进行面积、造价、合规性、能耗等关键数据的模拟分析,为建筑设计的多目标优化复杂决策提供全流程的决策支持;优化知识生产流程和知识分享方式,提高建筑行业整体专业水平和知识传播速度等。

建立垂直领域知识库、实现搜索优化与知识推动、自动知识生成与辅助决策,有助于推动传统建筑行业向数字化、智能化转型,也是建筑行业形成新质生产的关键步骤,期待更多同行的关注和研究。

参考文献:

- [1] 冯新翎,何胜,熊太纯,等.“科学知识图谱”与“Google 知识图谱”比较分析——基于知识管理理论视角[J]. 情报杂志,2017,36(1):149-153.
- [2] 黄恒琪,于娟,廖晓,等. 知识图谱研究综述[J]. 计算机系统应用,2019,28(6):1-12.
- [3] 阮彤,王梦婕,王昊奋,等. 垂直知识图谱的构建与应用研究[J]. 知识管理论坛,2016,1(3):226-234.
- [4] 陈翠,李鹏昊. 历史建筑商业化利用科学知识图谱分析与思
- [5] 江嘉玮. 学派与学术社会网络:一份刻画岭南建筑的知识图谱[J]. 建筑学报,2023(12):98-102.
- [6] JOHN HANCOCK CALLENDER. Time-saver standards for architectural design data[M]. New York: McGraw-Hill, 1982.
- [7] 杨玉基,许斌,胡家威,等. 一种准确而高效的领域知识图谱构建方法[J]. 软件学报,2018,29(10):2931-2947.
- [8] YANG Y J, XU B, HU J W, et al. Accurate and efficient method for constructing domain knowledge graph[J]. Journal of software, 2018, 29(10): 2931-2947.
- [9] 张博尧,曹荣强,万萌,等. 垂直领域知识图谱构建及应用平台的设计与实现[J]. 数据与计算发展前沿,2023,5(3):111-122.
- [10] ZHANG B Y, CAO R Q, WAN M, et al. Design and implementation of a platform for domain knowledge graph construction and applications[J]. Frontiers of data & computing, 2023, 5(3): 111-122.
- [11] 张兆锋. 基于知识图谱的技术功效图自动构建及其应用研究[D]. 南京:南京大学, 2018.
- [12] ZHANG Z F. Research on automatic construction and application of technology-efficiency map based on knowledge graph [D]. Nanjing: Nanjing University, 2018.
- [13] 彭鑫. 基于知识管理的企业知识图谱构建研究[D]. 武汉:武汉大学, 2018.
- [14] PENG X. The research on enterprise knowledge graph construction based on enterprise knowledge management [D]. Wuhan: Wuhan University, 2018.