

# 《标准知识图谱 第1部分：实现指南》（征求意见稿）

## 编制说明

### 一、工作简况

#### （一）任务来源

《标准知识图谱 第1部分：实现指南》是根据国家标准化管理委员会关于下达2024年第二批推荐性国家标准计划及相关标准外文版计划制定的，项目计划编号为“20240876-T-469”，计划完成时间为2025年。该计划项目由全国标准数字化标准化工作组（SAC/SWG 29）提出并归口。

#### （二）标准制定背景

2021年12月，国家标准委、中央网信办等部门联合印发《“十四五”推动高质量发展的国家标准体系建设规划》，明确提出要深入推进国家标准数字化试点，开展机器可读标准、开源标准、标准数字化平台标准等国家标准的研制工作。目前我国标准数字化工作尚处在纸质向电子化转换的初始阶段，相关标准研制尚未成体系，相关标准化管理制度和管理机构尚未配套，相关标准数字化平台尚未建设，因此，有必要推动相关标准化工作的开展，满足国家及社会对标准数字化的迫切需求。

当前由于标准的数据服务、信息服务仍处于初级阶段，大部分标准均采用纸质或简单电子版形态展现，机器难以读取、理解与使用，阻碍了标准数字化进程，从而影响智能服务应用技术的发展。标准知识图谱是标准数字化过程中的关键技术，是使用人工智能算法对标准及其生命周期全过程赋能，对标准所承载的规则知识建立完整且符合逻辑的关联体系，具备可通过数字设备读取、传输、理解以及使用的能力。标准知识图谱采用当前主流的先进知识图谱构建技术，能够对半结构化和非结构化的标准文本数据进行知识抽取、知识融合和知识表示，构建成为机器可读的形式参与计算，以实现知识检索、知识推理等功能，具有一定的先进性和创新性。本标准的制定，能有效促进标准数字化进程，提高标准知识高效利用，有助于标准知识图谱质量效益提升，提高标准知识图谱的使用价值和效率。

#### （三）起草过程

该项目下达后，由中国电子技术标准化研究院和广州赛西标准检测研究院有限公司共同组织起草，于2024年4月至6月期间，广泛征集参编单位，组成涵

盖产、学、研、用等相关单位的标准编制组，制定了标准编制工作计划，同步推进标准编制工作。

起草阶段工作如下：

2024年5月9日，编制组在北京召开了标准编制的启动会，来自标准知识图谱领域相关科研院所、高校和企业39家单位的40余名代表参加会议。与会专家就标准知识图谱的构建、更新和应用进行了详细的讨论，逐一分析了标准知识本体模型的构建、融合方法、更新方法，并对标准知识图谱的技术要求，标准知识图谱应用服务内容，标准知识图谱构建和应用相关示例进行了补充和完善；

2024年5月16日至2024年6月21日，中国电子技术标准化研究院和广州赛西标准检测研究院有限公司以线上线下相结合的形式，共同组织开展了6次研讨会，分别由南方电网科学研究院有限责任公司（广州）、卡斯柯信号有限公司（上海）、南京理工大学（南京）、比亚迪股份有限公司（深圳）、同方知网数字出版技术股份有限公司（北京）和中国汽车工程研究院股份有限公司（重庆）承办，来自科研院所、高校、行业协会、企业92家单位的200余位专家参加了会议。与会专家深入讨论了标准草案的适用范围、实现框架、构建与更新方法、质量评估、示例和涉及专利等内容，通过6次线上线下相结合研讨和现场编制，确立了由标准知识图谱实现框架、标准数据预处理，标准知识图谱构建、更新、评估和应用构成的标准框架及其内容，经反复修改和完善，形成标准征求意见稿和编制说明。

## 二、国家标准编制原则、主要内容和确定依据

### （一）国家标准编制原则

本标准按照 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》规定的文字表述和基本格式编写，编制程序符合相关文件的规定。

标准编制遵循“先进性、通用性、实用性、代表性”的原则，标准在编制过程中综合各行业标准知识图谱构建最新研究成果，归纳具有共性的实现方法和具有代表性的应用场景，形成完善的标准知识图谱实现指南，使其在保证先进性的同时具有广泛的普适性和可操作性，能为各行业构建、更新、评估和应用标准知识图谱提供指导。

## （二）标准的主要内容

本标准描述了标准数据预处理、标准知识图谱构建、更新和评估方法，提供了标准知识图谱应用建议，给出了标准知识图谱构建和行业应用的参考案例。

本标准适用于标准知识图谱的构建、评估、更新和应用。

## （三）标准主要内容的确定依据

本标准主要内容包括术语和定义、标准知识图谱实现框架、标准数据预处理、标准知识图谱构建与更新、标准知识图谱评估、标准知识图谱应用和附录，各部分内容的确定依据如下：

**术语和定义：**主要引用了 GB/T 42131—2022《人工智能 知识图谱技术框架》、GB/T 41867—2020《信息技术 人工智能 术语》、GB/T 36073—2018《数据管理能力成熟度评估模型》和 GB/T 32392.4—2015《信息技术 互操作性元模型框架(MFI) 第4部分：模型映射元模型》规定的术语和定义，并参照 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》，定义了“标准要素”“标准知识元素”“标准知识图谱”3项术语；

**标准知识图谱技术框架：**本章基于标准预处理、标准知识图谱构建与更新、标准知识图谱评估、标准知识图谱应用的内容及其相互关系进行归纳和凝练，形成标准知识图谱技术框架；

**标准数据预处理：**参照 GB/T 28448—2019《信息安全技术网络安全等级保护测评要求》、GB/T 31500—2015《信息安全技术存储介质数据恢复服务要求》、GB/T 22373—2021《标准文献元数据》、QB/CAII 04-03—2023《国家工业互联网大数据中心体系标准》等标准，结合行业内常用的数据预处理方法和实践经验制定；

**标准知识图谱构建与更新：**参照 GB/T 42131—2022《人工智能 知识图谱技术框架》、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》等标准，结合书籍《锻标成金—标准数字化知识图谱发展之路》以及相关制造业企业、院校和科研院所的研究成果和实践经验制定，构建与更新方法已经过多个科研项目和企业内部项目的验证。其中标准知识图谱本体构建与更新流程（图2）、本体建模方法（图3、图4和图5）主要参考了参考斯坦福本体构建七步法（Noy N F, McGuinness D L. Ontology development 101: A guide to creating

your first ontology[J]. 2001.), 经过标准知识图谱构建专家对产品、工业生产过程等标准数据结构化过程实践验证, 上述建模流程、方法能满足多种标准知识图谱构建场景的需求, 本体构建结果能准确覆盖标准要素, 为当前行业内最为广泛使用的标准知识图谱构建方法;

标准知识图谱评估: 基于 GB/T 42131—2022《人工智能 知识图谱技术框架》, 根据各行业标准知识图谱构建方和应用方的需求, 制定了标准知识图谱构建各环节的评估要点和评估方法, 评估指标与 GB/T 42131—2022 协调一致;

标准知识图谱应用: 从标准管理、研制和应用的角度出发, 根据各行业标准知识图谱的实际应用情况和先进应用案例, 提炼出标准知识图谱应用的基础功能要求和应用范围;

附录 A 和附录 B: 根据行业内已有的应用案例情况, 选取具有代表性的领域, 形成标准知识图谱构建示例和具体应用案例。

### **三、试验验证的分析、综述报告, 技术经济论证, 预期的经济效益、社会效益和生态效益**

#### **(1) 试验验证分析**

本标准主要技术内容包括标准数据预处理、标准知识图谱构建与更新、标准知识图谱评估、标准知识图谱应用及构建示例和应用案例, 主要依据现有相关标准和前沿实践经验制定, 标准知识图谱的构建方法已得到广泛应用, 技术内容得到各行业的实践验证, 并在编制组内开展了行业知识图谱的构建试验验证, 具有先进性、普适性和可操作性。

#### **(2) 预期效益**

标准知识图谱的需求源于标准数字化, 其产业化应用生态表现为标准数字化应用, 以及标准数字化创新, 通过对标准文本的数字化转化, 可对标准信息资源进行深层次加工, 充分挖掘和提升标准文献资源内在附加值。借助标准文本数字化转化的成果, 可以进行标准立项前的查新, 轻松、高效地进行标准之间在线技术指标的比对, 实现全生命周期、全流程跟踪管理及应用; 可通过对标准文本中的图形和表格进行检索和结果数据定向抽取, 实现了用户对所需标准信息的智能化检索, 保障从传统的标准信息模糊检索向高精度检索、文本挖掘、知识发现等方向转变。

本标准的研制符合国际标准化趋势和产业发展需求，一是助推标准数字化转型，有助于解决当前标准的数据服务、信息服务仍处于初级阶段，大部分标准均采用纸质或简单电子版形态展现，以及机器难以读取、理解与使用，阻碍标准数字化进程，影响智能服务应用技术发展的难题。二是抢占标准数字化制高点，在 ISO、CEN-CEELEC、ANSI 等标准化组织的标准化战略中均将标准数字化作为重要目标与布局领域，并已开展相关研究的战略部署。三是赋能标准知识图谱，理解标准知识图谱是标准数字化过程中的关键技术，是使用人工智能算法对标准及其生命周期全过程赋能。

#### 四、与国际、国外同类标准水平的对比情况

本标准制定过程中未查到同类国际、国外标准。

本标准水平为国内先进水平。

#### 五、以国际标准为基础的起草情况，以及是否合规引用或采用国际国外标准，并说明未采用国际标准的原因

本标准未采用国际标准或国外先进标准。

#### 六、与有关的现行法律、法规和标准的关系

本标准遵循有关产业政策，符合国家相关法律、法规的规定，隶属于标准数字化（SWG29）标准体系“B 过程实现—B.2 内容加工与处理”板块，标准的对象为标准知识图谱的构建、更新、评估、应用整个实现过程，主要解决如何抽取标准所承载的规则知识并建立完整且符合逻辑的关联体系，使其具备可通过数字设备读取、传输、理解以及使用能力的问题，为标准知识图谱的构建和实施提供指导和规范，确保图谱的准确性、一致性和可用性，支撑标准制定、管理、实施，是实现标准数字化的关键技术手段。

目前，SWG29 其他在研的标准包括：

- 1、《标准数字化 第 1 部分：通用指南》；
- 2、《标准机器语言表达 第 1 部分：基本架构与要素表达通用要求》；
- 3、《标准机器语言表达 第 2 部分：基于 XML 的标准文档内容标记指南》；
- 4、《标准数字化平台 第 1 部分：系统架构》；
- 5、《数字标准 标准信息模型架构》；

以上标准主要提出标准数字化活动、标准机器语言表达、标准数字化平台建

设、数字标准模型架构等方面的通用规则，标准知识图谱主要是建立这些规则底层标准知识间的关系，与标准体系中其他标准协调统一，可配套使用。本标准与 GB/T 42131—2022《人工智能 知识图谱技术框架》、YD/T 4044—2022《基于人工智能的知识图谱构建技术要求》等同类标准，在知识图谱的通用构建方法上保持协调一致，并针对标准的数据、结构特点、应用进行扩展和补充。

## 七、重大分歧意见的处理经过和依据

暂无。

## 八、涉及专利的有关说明

本标准在编制过程各种，识别到 7.2 与本体构建、标准知识获取、标准知识融合和标准知识存储可能涉及到相关的专利的使用。专利号及其专利名称如下：

序号	专利号	专利名称
1	ZL 2021 1 0733216.9	标准知识图谱构建、标准查询方法及装置
2	ZL 2022 1 0148740.4	多层结构标准知识图谱构建、标准检索方法及装置
3	ZL 2021 1 0337471.1	标准知识图谱构建、标准查询方法及装置
4	ZL 2021 1 0337459.0	标准文本标注、标准图谱构建方法及装置
5	ZL 2021 1 0588256.9	标准文本纠错方法、装置、电子设备和存储介质

上述专利实施许可声明方式：a)免费许可实施。

该专利持有人已向本文件的发布机构承诺，他愿意同任何申请人在合理且无歧视的条款和条件下。就专利授权许可进行谈判。该专利持有人的声明已在本文件的发布机构备案。相关信息可以通过以下联系方式获得：

专利持有人：中国电子技术标准化研究院/北京赛西科技发展有限责任公司。

地址：北京东城区安定门东大街 1 号。

## 九、实施国家标准的要求，以及组织措施、技术措施、过渡期和实施日期的建议等措施建议

建议本标准发布后 3 个月实施，通过标准宣贯会、研讨会等方式在知识图谱领域的相关研制单位、用户单位、科研机构等进行广泛的宣传和推广应用。

## 十、其他应予说明的事项

（项目延期/项目名称变更/采标类型变更/采标标准中内容的删除等）

无。

推荐性国家标准《标准知识图谱 第1部分：实现指南》

编制工作组

2024年8月7日