



中华人民共和国国家标准

GB/T 45257—2025

新闻出版 知识服务 知识元提取与标引

Press and publication—Knowledge services—
Extraction and marking up of knowledge element

2025-02-28 发布

2025-06-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	Ⅲ
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 提取与标引对象	2
5.1 提取对象	2
5.2 标引对象	2
6 提取与标引方式	2
6.1 提取原则与方式	2
6.2 标引原则与方式	3
7 提取规则与规格	3
7.1 提取规则	3
7.2 提取规格	5
8 标引规则与规格	6
8.1 标引用词	6
8.2 定位信息	6
8.3 标引定位	6
8.4 标引粒度	7
8.5 标引规则	7
8.6 标引规格	8
9 提取与标引流程	9
9.1 基本流程	9
9.2 提取流程描述	9
9.3 标引流程描述	10
10 提取与标引质量	10
10.1 提取基本要求	10
10.2 标引基本要求	10
参考文献	12

前 言

本文件按照 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》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国新闻出版标准化技术委员会（SAC/TC 527）归口。

本文件起草单位：国防工业出版社、时代出版传媒股份有限公司、中国建筑出版传媒有限公司、人教数字出版有限公司、北京拓标卓越信息技术研究院、中国新闻出版研究院、北京中企智造科技有限公司、保定市魁信达数据处理有限公司、同方知网数字出版技术股份有限公司、内蒙古出版集团有限责任公司、中国农业出版社有限公司、中国中医药出版社有限公司、河南教育电子音像出版社有限责任公司、天津大学出版社有限责任公司、研究出版社有限公司、北京航空航天大学、英大传媒投资集团有限公司。

本文件主要起草人：欧阳黎明、李明、卫敏、龙正武、邢剑飞、杜香和、周长岭、安秀敏、谢冰、刘颖丽、余敬春、袁华莉、杨继全、朱骞、彭静、马亮、王金波、布日古都、张瑶、孙海洋、苏合伟、王碑、贾慧君、张宝昌。

新闻出版 知识服务

知识元提取与标引

1 范围

本文件规定了知识元的提取与标引的对象、方式、规则、流程与质量要求。
本文件适用于新闻出版领域知识库和知识资源库的建设与管理。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 45256 新闻出版 知识服务 知识本体构建流程

3 术语和定义

下列术语和定义适用于本文件。

3.1

标引 marking-up

根据标引语言的规则，用词或短语对资源内容或形式表示的过程。

[来源：CY/T 101.1—2014，6.2.7，有修改]

3.2

版式文件 fixed-layout document

排版后生成的，包含版面固化呈现需要的全部数据的一种文件。

[来源：CY/T 101.1—2014，6.3.10]

3.3

流式文件 stream document

按照内容逻辑顺序，内容呈现可适应终端设备屏幕或窗口变化的一种文件。

[来源：CY/T 101.1—2014，6.3.11]

3.4

语素 morpheme

语言中最小的音义结合体。

3.5

知识元 knowledge element

在应用需求下，表达一个完整事物或概念的不必再分的独立的知识单位。

[来源：GB/T 38377—2019，2.3]

4 缩略语

ePub：电子图书标准（Electronic Publication）