

Q/HD

同方知网数字出版技术股份有限
公司企业标准

Q/HDZWS0005—2015

文献、知识元的元数据结构与
著录标准

2015—12—03 发布

2015—12—03 实施

同方知网数字出版技术股份有限公司 发布

目录

目录

目录	2
表目录	5
1 前言	1
2 范围	1
3 规范性引用文件	1
4 专业术语	1
4.1 CNKI 基础元数据	1
4.2 检索元数据	2
4.3 元数据项（元数据元素）	2
4.4 著录项错误率	2
4.5 著录文字错误率	2
4.6 标引正确率	2
4.7 标引一致度	2
4.8 统计资料	3
4.9 统计年鉴	3
4.10 年鉴来源	3
4.11 年鉴入编	3
4.12 稿件收录	3
4.13 统计资料类型	3
4.14 统计年鉴种数完整率	3
4.15 著录项错误率	3
4.16 著录文字错误率	4
4.17 著录内容版式错误率	4
5 应用原则	4
5.1 规范使用主体	4
6 元数据框架	4
6.1 概述	4
6.2 基础元数据框架	5
6.3 元数据扩展方式	5
6.3.1 上下位元数据元素扩展	5
6.3.2 并列元数据元素扩展	5
6.3.3 继承元数据元素扩展	5
6.3.4 元数据引用扩展	6
6.4 检索元数据选取原则	6
6.5 多分类体系	6
7 年鉴元数据与著录标准	6
7.1 收录原则与范围	6
7.1.1 收录原则	6
7.1.2 收录范围	7
7.1.3 收录完整性质量指标	7

7.1.4	收录及时性指标.....	8
7.2	多目元数据.....	8
7.2.1	多目元数据及其定义.....	8
7.2.2	多目元数据的规范控制.....	10
7.3	知识元元数据.....	10
7.3.1	数值知识元结构.....	10
7.3.2	统计数值知识元质量标准.....	10
7.4	元数据著录规则及其质量要求.....	11
7.4.1	著录元数据.....	11
7.4.2	标引元数据.....	12
8	期刊元数据与著录标准.....	13
8.1	元数据.....	13
8.1.1	元数据及对应关系.....	13
8.1.2	元数据的规范控制.....	15
8.2	元数据著录规则及其质量要求.....	15
8.2.1	著录元数据.....	15
8.2.2	标引元数据.....	16
9	工具书元数据与著录标准.....	17
9.1	元数据结构.....	17
9.1.1	元数据项定义.....	17
9.2	数据著录规则及其质量要求.....	19
9.2.1	著录数据.....	19
10	图片知识元数据标准.....	20
10.1	数据与定义.....	20
10.2	数据收录标准.....	22
10.3	各元数据质量标准.....	23
10.4	内容分类标准与类别定义.....	24
10.4.1	图片内容分类体系构建说明.....	24
10.4.2	CNKI 图片知识库图片分类原则.....	24
10.4.3	CNKI 图片知识库图片内容分类标准与类型定义.....	25
11	专家库数据规范.....	30
11.1	专家库元数据.....	30
11.1.1	元数据及其定义.....	30
11.1.2	元数据的规范控制.....	32
11.2	专家知网节的内容.....	33
11.2.1	节点专家.....	33
11.2.2	节点专家的相关专家连接.....	35
11.3	专家收录的完整性指标.....	36
11.3.1	专家库完整性质量指标.....	36
11.3.2	专家信息收集完整率.....	36
12	作者规范.....	36
12.1	作者姓名.....	36
12.1.1	作者姓名的选择.....	37
12.1.2	作者名称的书写.....	37

12.2	性别.....	37
12.2.1	简历.....	37
13	引文数据库规范.....	37
13.1	元数据.....	37
13.1.1	元数据及其定义.....	37
13.1.2	元数据的规范控制.....	39
13.2	元数据著录规则及其质量要求.....	39
13.3	引文数据库规范文档要求.....	40
13.3.1	参考文献规范.....	40
13.3.2	期刊名称规范.....	40
13.4	资源收录原则与规范.....	40
13.4.1	已收录全文或题录资源的参考文献部分的收录原则.....	41
13.4.2	未收录资源的参考文献部分的收录原则.....	41

表目录

表 6-1 CNKI 统一元数据框架与 DC 对映表.....	5
表 7-1 《中国经济社会发展统计数据库》元数据及其定义.....	8
表 7-2 数值知识元结构.....	10
表 7-3 数值知识元质量标准.....	10
表 7-4 元数据著录错误率质量指标.....	12
表 7-5 元数据标引质量指标.....	13
表 8-1 《中国学术文献网络出版总库》及各子库元数据对映表.....	13
表 8-2 元数据著录错误率质量指标.....	16
表 8-3 元数据标引质量指标.....	17
表 9-1 工具书元数据项及其定义.....	17
表 9-2 著录错误率质量指标.....	19
表 10-1 数据字段定义.....	21
表 10-2 元数据标引质量指标.....	23
表 10-3 标引数据标引质量指标.....	24
表 10-4 全文数据质量指标.....	24
表 11-1 专家库元数据及其定义.....	30
表 11-2 元数据标引质量指标.....	32
表 13-1 《中国引文数据库》元数据及其定义.....	37

1 前言

本标准由同方知网数字出版技术股份有限公司提出；
 本标准起草单位：同方知网数字出版技术股份有限公司；
 本标准主要起草人：王明亮、张振海、熊海涛、梁洵、汪新红、丁慎训、万锦堃、李小红、欧坤、王国红、赵纪元、师庆辉、陈华、冯自强、康欢；
 本标准于 2015 年 12 月首次发布。

2 范围

本标准规范了 CNKI 元数据描述框架，以及元数据结构及其数据著录标准要求；
 本标准适用于学习需求驱动下的数字出版资源定制投送系统及应用示范项目。

3 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。

凡是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。

Q/CNKI 125.1-2012	《中国学术文献网络出版总库》
Q/CNKI 501	CNKI 数字对象命名规范
ISO/IEC 11179	数据元素规范描述和标准化（Specification and Standardization of Data Elements）
ISO 15836-2003	信息和文献工作.都柏林核心元数据元素集（Information and Documentation — The Dublin Core Metadata Element Set ）
Q/CNKI 201.4-2010	《中国工具书网络出版总库》第四部分元数据结构及数据质量
Q/CNKI 202-2012	《中国经济社会发展统计数据库》产品标准
	图片知识元数据库产品标准

4 专业术语

4.1 CNKI 基础元数据

CNKI 系列数据库元数据所采用的、作为所有元数据基础的 15 项元数据：题名、作者、主题、描述、出版者、时间特征、资源类型、格式、标识符、来源、语种、关联、覆盖范围、权限、全文文本。参照都柏林核心元数据集（Dublin Core，以下简称 DC）规范制定。

4.2 检索元数据

直接用于提供检索功能的元数据。

4.3 元数据项（元数据元素）

对数据对象某一特征的描述，构成元数据记录的基本单位。

4.4 著录项错误率

元数据项著录错误的元数据项数量与元数据项总数的比率。

$\text{著录项错误率} = (\text{元数据项著录错误的元数据项数量} / \text{元数据项数量}) \times 100\%$

著录项错误是按元数据项计算，凡未能对元数据项进行客观著录的情况均计为著录项错误。适用于文字错误对元数据项内容影响大的元数据项。

元数据项内容为单值的，计为一项；元数据项内容为多值的，则按值的数量计算项数，例如，一篇文献有4个关键词，则计为4项。

4.5 著录文字错误率

元数据项内容文字错误的数量与元数据项文字总数的比率。

$\text{著录文字错误率} = (\text{元数据项内容文字错误的数量} / \text{元数据项文字数量}) \times 100\%$

著录文字错误按文字计算，适用于文字错误对元数据内容影响较小的元数据。元数据内容中文字与原文不符的计为著录文字错误，但不包括对原文的文字错误进行纠正的情况。

4.6 标引正确率

元数据标引正确的元数据项数量和元数据项总数的比率。

$\text{标引正确率} = (\text{元数据标引正确的元数据项数量} / \text{元数据项数量}) \times 100\%$

标引项目基本反映文献主题，无原则性错误，则计为标引正确。

4.7 标引一致度

元数据标引一致的元数据项数量与元数据项总数的比率。

$\text{标引一致度} = (\text{元数据标引一致的元数据项数量} / \text{元数据项数量}) \times 100\%$

标引一致度反映同主题文献标引一致的程度。同主题文献，被标引的主题类元数据应保持一致。

4.8 统计资料

统计活动过程中所产生的、反映国民经济和社会发展情况的统计成果以及与之关联的其他资料的总称。

4.9 统计年鉴

统计资料逐年更新统计数字而形成的年度出版物。

4.10 年鉴来源

依约向《中国经济社会发展统计数据库》提供年鉴文献资料及其著作权使用许可的年鉴编辑部或杂志社，下统称为年鉴编辑部。

4.11 年鉴入编

与年鉴编辑部签订协议后，该年鉴正式入编。

4.12 稿件收录

根据年鉴入编协议，通过与年鉴编辑部联系、图书馆借阅、自行购买等渠道取回稿件。

4.13 统计资料类型

统计年鉴是统计资料的主要类型，按照统计资料自身的分类，把统计资料整刊分为统计年鉴、统计摘要、普查资料、调查资料、历史统计资料汇编、分析报告共计六类。

4.14 统计年鉴种数完整率

《中国经济社会发展统计数据库》实际收录的统计年鉴种数与《中国年鉴概览》中和其他已知的和已停止出版的统计年鉴中符合《中国经济社会发展统计数据库》收录原则的统计年鉴种数之比。

4.15 著录项错误率

著录错误的数据项数量与数据项总数的比率。

$$\text{著录项错误率} = (\text{著录错误的数据项数量} / \text{数据项数量}) \times 100\%$$

著录项错误是凡未能对数据项进行客观著录的情况均计为著录项错误。

数据项内容为单值的，计为一项；数据项内容为多值的，则按值的数量计算项数，例如，

一册工具书有 4 个主要责任者，则计为 4 项。

4.16 著录文字错误率

数据项内容文字错误的数量与数据项文字总数的比率。

$$\text{著录文字错误率} = (\text{数据项内容文字错误的数量} / \text{数据项文字数量}) \times 100\%$$

著录文字错误按文字计算，适用于文字错误对数据内容影响较小的元数据。数据中文字与原文不符的计为著录文字错误，但不包括对原文的文字错误进行纠正的情况。

4.17 著录内容版式错误率

数据项文本、表格、图形、图片、公式显示样式著录错误的的数据项数量和数据项数量的比率。

$$\text{著录内容版式错误率} = (\text{数据版式错误的的数据项数量} / \text{数据项数量}) \times 100\%$$

著录内容版式能正确反映印刷版工具书的原始含义，并能在全文显示中正确阅读，则计为著录正确。

5 应用原则

5.1 规范使用主体

规范使用主体包括：

- a) 同方知网数字出版技术股份有限公司；
- b) 社会科学文献出版社；
- c) 北京经济管理职业学院；
- d) 中国人民解放军总后勤部后勤科学研究所。

6 元数据框架

6.1 概述

CNKI 元数据框架是制定 CNKI 系列数据库元数据的基础。它是规范各数据库元数据、提供统一检索功能的需要，也是进行资源整合的基础。

CNKI 元数据框架参照 DC 制定，并根据标准和产品的特点及需要，对元数据元素的描述方法和元数据元素进行了调整和扩展。由于创建者与其他责任者之间的差别很小，普通使用者较难区别，为方便应用，CNKI 元数据框架将 DC 中创建者（Creator）与其他责任者（Contributor）项合并为作者项。

基于全文数据库产品的特点，CNKI 元数据元素扩充了全文元素，更好地描述和记录文献的全文文本。

6.2 基础元数据框架

CNKI 所有数据库的元数据均须依据基础元数据框架制定，复用 CNKI 基础元数据的元素。不能复用时，可按照扩展规则产生。表 6-1 列出了基础元数据框架中各项元素的名称，标签，定义以及对映的 DC 元数据元素。

表 6-1 CNKI 统一元数据框架与 DC 对映表

元素名称	标签	定义	对映的 DC 元数据元素
题名	题名	文献资源的篇名。	题名 (Title)
作者	作者	资源的创建者或与创建有关者。	创建者 (Creator), 其他责任者 (Contributor)
主题	主题	内容主题描述。	主题 (Subject)
描述	描述	资源的解释、有关说明与其他内容信息。	描述 (Description)
时间特征	时间特征	与时间有关的资源出版特征描述。如出版时间、杂志的年卷期等。	日期 (Date)
出版者	出版者	资源的出版者。	出版者 (Publisher)
格式	格式	资源的物理媒体特性或数字媒体特性。	格式 (Format)
标识符	标识符	资源的明确的唯一标识。	标识符 (Identifier)
来源	来源	资源的出处。	来源 (Source)
语种	语种	资源的语言种类。	语种 (Language)
关联	关联	相关资源参照。	关联 (Relation)
资源类型	资源类型	资源出版类型。	类型 (Type)
覆盖范围	覆盖范围	内容覆盖的空间或时间范围。	覆盖范围 (Coverage)
权限	权限	资源的版权特性或访问权限。	权限 (Rights)
全文	全文	全文文本	

6.3 元数据扩展方式

6.3.1 上下位元数据元素扩展

上位元数据元素涵盖下位元数据元素的描述内容。下位元数据元素是对上位元数据元素的精细描述。

6.3.2 并列元数据元素扩展

扩展的元数据元素之间相对独立。

6.3.3 继承元数据元素扩展

通过继承其他元数据元素的属性扩展元数据元素。既可以是上下位继承扩展，也可以是并列继承扩展。

6.3.4 元数据引用扩展

通过引用其他元数据元素的属性扩展元数据元素。

6.4 检索元数据选取原则

设计检索元数据考虑下列因素：

- a) 生产成本；
- b) 检索控制精度；
- c) 数据加工技术难度。

6.5 多分类体系

主题元数据应采用多个分类体系，可以从多个不同的角度揭示文献内容的内在关系。

7 年鉴元数据与著录标准

7.1 收录原则与范围

7.1.1 收录原则

7.1.1.1 择优收录原则

对于中央级出版社出版的统计年鉴和统计资料优先收录。

7.1.1.2 收录完整性

中央级、省级统计年鉴原则上收全，其他统计年鉴原则上择优收录。

7.1.1.3 合法出版原则与著作权使用许可

收录出版年鉴应符合国家出版管理的有关法规文件；所收录的统计年鉴文献应依法、依约取得著作权人的著作权使用许可，包括但不限于数字化复制权、汇编权、信息网络传播权、电子出版权等权力。

7.1.1.4 及时性收录原则

纸质统计年鉴出版之后，要及时采集和入编纸质版稿件及附属电子文件。

7.1.2 收录范围

7.1.2.1 统计年鉴

《中国经济社会发展统计数据库》的收录范围为年鉴、年刊和具有相似内容的政报、公报。

7.1.2.2 经济运行进度数据

国家及各省市统计局、其他政府统计部门历年及最新经济运行进度数据。

7.1.2.3 统计部门其他公开资料

包括国家统计局、各省级统计局等统计信息网；发改委、商务部、中国人民银行、海关总署等部委网站。

7.1.2.4 国际数据

国际统计年鉴、世界银行、OECD、联合国开发计划署等国际组织的数据。

7.1.3 收录完整性质量指标

统计年鉴文献资源收录完整性的质量指标。

7.1.3.1 收录完整性质量指标

《中国经济社会发展统计数据库》中收录的中央级年鉴、省级年鉴的刊种完整率不低于95%，其他年鉴的刊种完整率不低于90%。

7.1.3.2 文献收录期数完整率

文献收录期数完整率为《中国经济社会发展统计数据库》所收录的出版的刊物期数与同期实际出版的年鉴总期数之比；文献收录期数完整率不低于95%；数据库实际收录年鉴的文献总篇数与收录年鉴出版的文献总篇数之比；文献篇数收录完整率不低于99.9%。

7.1.3.3 全文信息完整率

全文信息完整率不低于 99.9%。

7.1.4 收录及时性指标

$$\text{稿件收录准期率} = \frac{\text{准期收录的稿件}}{\text{全部收录的稿件}} \times 100\%$$

7.2 多目元数据

7.2.1 多目元数据及其定义

《中国经济社会发展统计数据库》的条目元数据及其定义见表 7-1。

表 7-1 《中国经济社会发展统计数据库》元数据及其定义

元素名称	名称	标签	定义
题名	条目题名	篇名	文献的题名。包括正题名、副题名、并列题名等。
	其他语种题名	其他语种篇名	文献题名的其他语种描述。包括其他语言正题名、副题名、并列题名等。
作者	作者	作者	对文献实体或数据库记录负有责任的团体名称或个人名称。
	机构	机构	文献发表时作者所属机构的名称，或文献提供的公告单位，或文献单位的责任者
主题	中国图书馆分类法分类号	中国图书馆分类法分类号	按照中国图书馆分类法体系，根据文献内容所涉及的主题指定的类别号。
	文献标识码	条目类型	按照年鉴条目中经常出现的文献类型进行划分。目前划分为工作总结、领导讲话、工作计划、事实资料、法律法规、统计公报、工作研究、政府文件、标准、人物、科研论文、大事记、图片、机构介绍、作品以及其他十六类文献
描述	说明	说明	揭示文献内容主题的一段文字。如图片说明等
	地域	地域	揭示文献内容所方式地点、或与文献内容关系最密切的地域特征，根据中国行政区划划分。
	目录编码	目录编码	描述单篇文献在所在年鉴目录体系中所在位置的一组代码
日期	网络出版日期	网络出版日期	文献加工完成后出版到数据库中的日期。

元素名称	名称	标签	定义
	年鉴年份	年鉴年份	揭示年鉴条目描述时间截止下限的年份
出版者	出版单位	出版单位	使印刷版文献成为可取得或利用状态的责任者。
	出版地	出版地	印刷版出版者注册所在地，根据中国行政区划划分。
格式	页数	页数	文献原文在其原始来源中所占用的页数。
	页码	页码	文献原文在其原始来源中，每页的编号或所有页的编号范围，或文献在其原始来源中的版号。
	原文文件大小	原文文件大小	文献正文的数字化原文的计算机文件的字节数。
标识符	文献统一编号	文献号	根据《中国学术期刊(光盘版)检索与评价数据规范》(CAJ-CD B/T 1-1998)中第5章关于文献编号的规定，为文献指定的、全世界范围内唯一的文献编号。
来源	年鉴中文名	年鉴中文名	中文版年鉴的正式名称，或外文版年鉴的正式中文翻译名称。对应于“来源名称”。
	年鉴英文名	年鉴英文名	英文版年鉴的正式名称，或非英文版年鉴的正式英文翻译名称。
	出版年卷期号	出版年卷期号	年鉴的年份、卷号及期号。
	国际标准连续出版物号	ISSN	根据国际标准组织 1975 年制订之 ISO-3297 的规定，由设于法国巴黎的国际期刊资料系统中心 (International Serial Data System – ISDS International Centre) 所赋予，申请登记的具有唯一性、可识别性、国际通行性的统一编号。
	中国出版物编号	CN	中国出版物标准编号。
	电子版国际标准连续出版物号	ISSN	根据国际标准组织 1975 年制订之 ISO-3297 的规定，由设于法国巴黎的国际期刊资料系统中心 (International Serial Data System – ISDS International Centre) 所赋予申请登记的具有唯一性、可识别性、国际性的统一编号。
	中国电子出版物编号	CN	中国电子出版物标准编号。
	电子版年期号	电子版年期号	以光盘形式发布时，电子文件所在光盘的编号。
语种	语种	语种	文献所使用的主要语言。
关联	参考文献	参考文献	以文后参考书目或脚注的形式列出的文献题录。

7.2.2 多目元数据的规范控制

对表 7-1 中的部分元数据使用适当的规范文档，以便对其检索功能进行规范控制。可以根据具体情况选择前控或后控方式。

- a) 对于作者，使用作者规范文档进行规范控制；
- b) 对于年鉴名称，使用年鉴名称规范文档进行规范控制；
- c) 对于机构，使用机构规范文档进行规范控制。

7.3 知识元元数据

7.3.1 数值知识元结构

《中国经济社会发展统计数据库》的数值知识元是一个六元组（时间、主体、指标、谓词、数值、单位）。如表 7-2 所示。

表 7-2 数值知识元结构

时间	主体	指标	谓词	数值	单位
2005 年	盘龙区	工业总产值	完成	62.93	亿元

7.3.2 统计数值知识元质量标准

《中国经济社会发展统计数据库》的数值知识元的质量标准为错误率 $e \leq 1\%$ ，检测方法采用分层随机抽样法，抽样结果采用假设检验法进行分析。

表 7-3 数值知识元质量标准

数值知识元_指标名称	<0.1%
数值知识元_数值	<0.01%
数值知识元_单位	<0.1%
数值知识元_其他要素	<0.1%

设当前数值知识元库中的统计年鉴本数为 I ，数值知识元总体数 $N = \sum_{i=1}^I N_i$ ， N_i 为第 i 本统计年鉴对应的数值知识元总数，其中 $1 \leq i \leq I$ 。现对每一种统计年鉴采用如下分层随机抽样：对每一本统计年鉴 $E_i (1 \leq i \leq I)$ ，抽样的表格数为该本内所有表格数的 $1/3$ ，对应每一表格中抽样的数值知识元为 10 个，并保证最终所抽样的数值知识元数量 $M_i \approx 0.01 \times N_i$ 。

采用假设检验法对 $E_i (1 \leq i \leq I)$ 的抽样数据进行检测，步骤如下：

- ① 建立假设，确定显著性 α 水平

$$H_0: e=0.001, H_1: e>0.001, \alpha=0.01(\text{置信度为 } 99\%)$$

- ② 计算统计量

$$Z_i = \frac{\mu_n - np_0}{\sqrt{np_0(1-p_0)}} = \frac{e_i - M_i \times 0.001}{\sqrt{M_i \times 0.001(1-0.001)}}$$

其中, e_i 为 E_i 中检测出的错误数, M_i 为统计年鉴 E_i 的抽样数。

③对于给定的显著性水平 $\alpha=0.01$, 其临界值为 $z_{0.99}=2.38$, 确定临界域为 $[2.38, +\infty)$ 。

④如果 $Z_i < z_{0.99}$, 则接受假设 H_0 , 即认为 E_i 的数值知识元抽取结果符合质量标准;

如果 $Z_i > z_{0.99}$, 则接受假设 H_1 , 即认为 E_i 的数值知识元抽取结果不符合质量标准。

根据每一本统计年鉴的检测结果, 可计算当前数值知识元库的质量水平:

①建立假设, 确定显著性水平

$H_0: e=0.001, H_1: e>0.001, \alpha=0.01$ (置信度为 99%)

②计算统计量

$$Z = \frac{\mu_n - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\sum_{i=1}^I e_i - \sum_{i=1}^I M_i \times 0.001}{\sqrt{\sum_{i=1}^I M_i \times 0.001(1-0.001)}}$$

其中, e_i 为 E_i 中检测出的错误数, M_i 为统计年鉴 E_i 的抽样数。

③对于给定的显著性水平 $\alpha=0.01$, 其临界值为 $z_{0.99}=2.38$, 确定临界域为 $[2.38, +\infty)$ 。

④如果 $Z < z_{0.99}$, 则接受假设 H_0 , 即认为当前数值知识元库中的质量符合标准; 如果

$Z > z_{0.99}$, 则接受假设 H_1 , 即认为当前数值知识元库中的质量不符合标准。

7.4 元数据著录规则及其质量要求

7.4.1 著录元数据

7.4.1.1 元数据著录

直接根据文献信息进行著录。

7.4.1.2 著录元数据元素

包括: 条目题名、其他语种题名、作者中文名、其他语种作者姓名、第一责任人、机构、

说明、页数、页码、目录编码、年鉴中文名、年鉴英文名、年鉴年份、卷期号、国际标准连续出版物号、中国出版物编号、语种、参考文献。

7.4.1.3 著录要求

元数据元素的著录应根据文献的相关信息和内容客观著录。

7.4.1.4 质量指标

著录元数据的质量指标包括两种类型：著录项错误率和著录文字错误率。

7.4.1.5 质量指标要求

各项元数据的质量指标及指标值见表 7-4。

表 7-4 元数据著录错误率质量指标

元数据名称	质量指标	指标值
条目题名	著录文字错误率	<0.05%
目录编码	著录项错误率	<0.05%
其他语种题名	著录文字错误率	<0.08%
作者	著录项错误率	<0.05%
机构	著录项错误率	<0.08%
说明	著录文字错误率	<0.1%
页数	著录项错误率	<0.05%
页码	著录项错误率	<0.05%
年鉴中文名、年鉴英文名	著录项错误率	<0.01%
年鉴年份	著录项错误率	<0.01%
出版年卷期号	著录项错误率	<0.01%
国际标准连续出版物号	著录项错误率	<0.01%
中国出版物编号	著录项错误率	<0.01%
语种	著录项错误率	<0.01%
参考文献	著录文字错误率	<0.1%

7.4.2 标引元数据

7.4.2.1 元数据标引

对每篇文献给出一些可以作为检索标识的元数据。

7.4.2.2 标引元数据项

标引元数据项包括：中国图书馆分类法分类号、年鉴条目类型、地域。

7.4.2.3 标引要求

- a) 准确、全面地进行标引；
- b) 中国图书馆分类法分类号按照《中国图书馆分类法》分类标引；
- c) 主题词按照《汉语主题词表》进行主题词标引。

7.4.2.4 质量指标

对于反映文献主题的标引元数据，有两个质量指标：标引正确率和标引一致度。

7.4.2.5 质量指标要求

各标引元数据的质量指标要求见表 7-5。

表 7-5 元数据标引质量指标

元数据名称	质量指标	质量指标要求
条目类型	标引正确率	≥90%
	标引一致度	≥90%
中国图书馆分类法分类号	标引正确率	≥90%
	标引一致度	≥90%
地域	标引正确率	≥98%
	标引一致度	≥98%

8 期刊元数据与著录标准

8.1 元数据

8.1.1 元数据及对应关系

《中国学术文献网络出版总库》的元数据及与各子库的元数据对应关系见表 8-1。

表 8-1 《中国学术文献网络出版总库》及各子库元数据对映表

元素名称	《总库》	期刊辑刊	学位论文	会议论文	报纸文献	标准	专利	科技成果
题名	题名	篇名 英文篇名	题名 英文题名 副题名	题名 英文题名	正标题 副标题 引题	标准号 中文名称 英文名称	专利名称	成果名称

			英文副题名					
作者	作者	中文作者名 英文作者名	作者 导师	中文作者名 英文作者名	作者	提出部门 起草单位 标准技术委员会 起草人	申请人 发明人	成果完成人
	第一责任人	第一作者	——	第一作者	第一责任人		——	——
	机构	作者单位	学位授予单位 作者单位 导师单位	作者单位	作者单位		申请机构	第一完成单位 合作完成单位
主题	关键词	中文关键词 英文关键词 机标关键词	中文关键词 英文关键词 机标关键词	中文关键词 英文关键词 机标关键词	机标关键词	关键词 中文主题词 英文主题词	机标关键词	关键词
	主题词	主题词	主题词	主题词	主题词	主题词	主题词	主题词
	中图分类号	中图分类号	中图分类号	中图分类号	中图分类号	中图分类号	中图分类号 专利分类号	中图分类号
描述	摘要	中文摘要 英文摘要	中文摘要 英文摘要 中文目录 英文目录	中文摘要 英文摘要	——	摘要	摘要 主权项	成果简介
	基金	基金	基金	基金	——	——	——	课题立项名称 课题立项编号
日期	发表时间	发表时间	发表时间 学位授予年度	发表时间 会议时间	——	发布日期 实施日期	申请日 公开日	研究起止时间 评价日期
	网络出版日期	网络出版日期	网络出版日期	网络出版日期	网络出版日期	网络出版日期	网络出版日期	网络出版日期
格式	页数	页数	页数	页数	页数	总页数	页数	
	页码	页码	页码	页码		——	——	——
	原文文件大小	原文文件大小	原文文件大小	原文文件大小	原文文件大小	——	——	——
标识	文献标识	文献标识	文献标识	文献标识	文献标	——	——	——

符	码	码	码	码	识码			
来源	文献来源	中文刊名 英文刊名 出版年卷 期号 期刊基本 参数 ISSN CN	学位授予 单位	中文会议 录名 称 英文会议 录名 称 会议录基 本参 数	报纸中 文名 报纸基 本参数	发布单位	申请人	成果完成 单位
关联	参考文献	参考文献	参考文献	参考文献	——	——	——	——
	引证文献	引证文献	引证文献	引证文献	引证文 献	引证文献	引证文献	——

8.1.2 元数据的规范控制

对表 8-1 中的部分元数据使用适当的规范文档，以便对其检索功能进行规范控制，可以根据具体情况选择前控方式或后控方式。

- a) 对于作者、第一责任人，使用作者规范文档进行规范控制；
- b) 对于期刊名称，使用期刊名称规范文档进行规范控制；
- c) 对于关键词、主题词，使用主题词表进行规范控制；
- d) 对于机构，使用机构规范文档进行规范控制。

8.2 元数据著录规则及其质量要求

8.2.1 著录元数据

8.2.1.1 元数据著录

直接根据文献信息进行著录。

8.2.1.2 著录元数据

著录元数据包括：条目题名、其他语种题名、作者中文名、其他语种作者姓名、第一责任人、机构、说明、页数、页码、目录编码、年鉴中文名、年鉴英文名、年鉴年份、卷期号、国际标准连续出版物号、中国出版物编号、语种、参考文献。

8.2.1.3 著录要求

元数据的著录应根据文献的相关信息和内容客观著录。

8.2.1.4 质量指标

著录元数据的质量指标包括两种类型：

- a) 著录项错误率；
- b) 著录文字错误率。

8.2.1.5 质量指标要求

各项元数据的质量指标及指标值见表 8-2。

表 8-2 元数据著录错误率质量指标

元数据名称	质量指标	指标值
中文题名	著录文字错误率	<0.01%
其他语种题名	著录文字错误率	<0.08%
作者中文名	著录项错误率	<0.01%
第一责任人	著录项错误率	<0.01%
作者机构	著录项错误率	<0.01%
作者基本信息	著录文字错误率	<0.08%
中文摘要	著录文字错误率	<0.01%
英文摘要	著录文字错误率	<0.1%
基金	著录文字错误率	<0.01%
页数	著录项错误率	<0.01%
页码	著录项错误率	<0.01%
出版来源	著录项错误率	<0.01%
参考文献	著录文字错误率	<0.1%

8.2.2 标引元数据

8.2.2.1 元数据标引

依据一定的规则，对每篇文献的内容特征给出一些可以作为检索标识的元数据。

8.2.2.2 标引元数据项

标引元数据项包括：主题词、中文关键词、英文关键词、中国图书馆分类法分类号、文献标识码。

8.2.2.3 标引要求

- a) 准确、全面地进行标引。
- b) 中国图书馆分类法分类号按照《中国图书馆分类法》（第五版）分类标引。

8.2.2.4 质量指标

对于反映文献主题标引元数据，有两个质量指标：标引正确率和标引一致度。

8.2.2.5 质量指标要求

各标引元数据的质量指标要求见表 8-3。

表 8-3 元数据标引质量指标

元数据名称	质量指标	质量指标要求
主题词	标引正确率	≥90%
	标引一致度	≥90%
中文关键词 英文关键词	标引正确率	≥90%
	标引一致度	≥90%
中国图书馆分类法分类号	标引正确率	≥90%
	标引一致度	≥90%

9 工具书元数据与著录标准

9.1 元数据结构

9.1.1 元数据项定义

工具书元数据项及其定义见表 9-1。

表 9-1 工具书元数据项及其定义

数据项名称	细分名称	定义
书名	书名	工具书的书名。包括书名、副书名、并列正书名、其他书名等。

数据项名称	细分名称	定义
	分辑名	丛书类工具书的分辑卷册的名称。
	丛书名	丛书类工具书的总卷名称。
责任者	主要责任者	对单册工具书负有责任的团体名称或个人名称。
	责任方式	单册书责任者对应的责任方式，包括编撰、编等
	丛书责任者	对丛书负有责任的团体名称或个人名称。
	丛书责任方式	丛书责任者对应的责任方式，包括编撰、编等。
来源	出版社	使印刷版文献成为可取得或利用状态的责任者。
	出版地	印刷版出版者注册所在地，根据中国行政区划划分。
	出版日期	工具书的出版日期，由年份和月份组成。
书号	ISBN 国际标准书号	根据国际标准组织 1975 年制订之 ISO-3297 的规定，由设于法国巴黎的国际期刊资料系统中心（International Serial Data System - ISDS International Centre）所赋予，申请登记的具有唯一性、可识别性、国际通行性的统一编号。
	统一书号	一般印在图书版权页和封底下端的、包括图书分类号、出版社代号和序号的组合数码。1956 年 2 月，中华人民共和国文化部出版事业管理局颁发《全国图书统一编号方案》（简称全国统一书号），并付诸实施。国家标准局于 1986 年 1 月批准颁布了中国标准书号，国家出版局决定，自 1987 年 1 月 1 日起，中国标准书号取代全国统一书号。
分类	整书中图分类号	中图分类号——是指采用《中国图书馆分类法》对科技文献进行主题分析，并依照文献内容的学科属性和特征，分门别类地组织文献，所获取的分类代号。
其他基本信息	版次	是用以统计版本内容的重要变更的。图书第一次出版为第一版，内容经过较大增删后出版为第二版，以下类推。同一种图书内容无变化，而只改换书名、开本、版式、装订、封面、出版者等，均不作为再版，即不作为版次的变更。
	装帧	书刊的结构与形态的设计，分为精装和简装两种。
	页码	单册工具书的总页数。
	字数	单册工具书的总字数。
	开本	印刷出版物成品幅面的大小。
	印数	单册工具书的总发行量。
	定价	单册工具书的纸质书定价。
	作品语种	工具书印刷发行的语种。
参考文献	参考文献	工具书的参考文献。

9.2 数据著录规则及其质量要求

9.2.1 著录数据

9.2.1.1 数据著录

直接根据文献信息进行著录。

9.2.1.2 著录数据元素

包括：书名、责任者、来源、书号、参考文献、条目题名、条目正文、条目作者、条目参考文献等。

9.2.1.3 著录要求

数据元素的著录应根据文献的相关信息和内容客观著录。

9.2.1.4 质量指标

著录元数据的质量指标包括两种类型：著录项错误率和著录文字错误率。

9.2.1.5 质量指标要求

各数据项的质量指标及指标值见表 9-2。

表 9-2 著录错误率质量指标

数据名称	质量指标	指标值
书名	著录项错误率	<0.01%
分辑名	著录项错误率	<0.01%
丛书名	著录项错误率	<0.01%
主要责任者	著录项错误率	<0.01%
责任方式	著录项错误率	<0.01%
丛书责任者	著录项错误率	<0.01%
丛书责任方式	著录项错误率	<0.01%
出版社	著录项错误率	<0.01%
出版地	著录项错误率	<0.01%
出版日期	著录项错误率	<0.01%

数据名称	质量指标	指标值
ISBN 国际标准书号	著录项错误率	<0.01%
统一书号	著录项错误率	<0.01%
整书中图分类号	著录项错误率	<0.01%
版次	著录项错误率	<0.05%
装帧	著录项错误率	<0.05%
页码	著录项错误率	<0.05%
字数	著录项错误率	<0.05%
开本	著录项错误率	<0.05%
印数	著录项错误率	<0.05%
定价	著录项错误率	<0.05%
作品语种	著录项错误率	<0.01%
参考文献	著录文字错误率	<0.01%
条目名称	著录文字错误率	<0.01%
条目的英文	著录文字错误率	<0.01%
条目拼音	著录文字错误率	<0.01%
词目异体	著录文字错误率	<0.01%
条目释文	著录文字错误率	<0.01%
	著录内容版式错误率	<0.01%
条目注释	著录文字错误率	<0.01%
条目的说明	著录文字错误率	<0.01%
条目作者	著录文字错误率	<0.01%
责任方式	著录文字错误率	<0.01%
条目参考文献	著录文字错误率	<0.01%

10 图片知识元数据标准

10.1 数据与定义

CNKI 图片知识库包含下列几类数据字段：

1. 图片基本信息字段：图片中文关键词、图片英文标题、图片说明、图片说明等；
2. 文献信息关联字段：文献篇名、作者，中文刊名，机构、专题代码、SYS-VSM、作者代码、机构代码、中文摘要，学科分类信息等；
3. 语义内容标引字段：图片中文关键词、图词、图片关键词 VSM、图片内容分类等；
4. 视觉特征：灰度、纹理、拓扑、骨架等图像视觉特征；
5. 关联推荐图片：读者推荐图片、语义相似图片、视觉相似图片、参考文献图片、引用文

献图片、同文图片、同作者图片等；
具体各字段定义如下：

表 10-1 数据字段定义

元素名称	名称	标签	定义
元数据	图片中文标题	图片中文标题	图片特征的中文描述。
	图片英文标题	图片英文标题	图片特征的英文描述。
	图片说明	图片说明	对图片内容的说明性描述。
	图片页码	图片页码	图片在文献中所处的位置。
标引数据	ID	图片 ID	人工标引的图片地址。
	图片关键词	图片关键词	描述图片内容的一个或多个规范词。
	图片关键词 VSM	图片关键词 VSM	描述图片内容的一个或多个规范词向量。
	图词	图词	描述图片视觉基本特征的关键词。
	一级分类	一级分类	将图片特征分为若干类，描述图片最基本特征的类别，即为一级分类。例如：曲线图中的曲线就是图片的最基本特征。
	二级分类	二级分类	一级分类的子类别。例如：心电图即为曲线图的二级分类。
	FEATURE	FEATURE	描述图片内容的特征。
	图片地址	图片地址	图片所在的数据库位置。
	图片尺寸	图片尺寸	图片的面积大小。
	图片大小	图片大小	图片所占的空间位置。
	图片梯度	图片梯度	图片清晰度的度量。
	同文图片 ID	同文图片 ID	在同一篇文章中其他图片的人工标引地址。
	语义相关图片 ID	语义相关图片 ID	与图片关键词相关的图片的人工标引地址。
	视觉相似图片 ID	视觉相似图片 ID	与图片在视觉上相似的图片的人工标引地址。
文献与知识 链接数据	篇名	篇名	图片所在文献的标题名称。
	作者	作者	对文献实体或数据库记录负有责任的团体或个人名称。
	中文刊名	中文刊名	中文版期刊的正式名称，或外文版期刊的正式中文翻译名称。
	拼音刊名	拼音刊名	中文版期刊正式名称的汉语拼音表述。
	年	年	文献出版的年份。
	期	期	文献出版的期号。
	文件名	文件名	文献出版时被赋予的出版分类编号。
	页	页	文献出版时所使用的页码。
	页数	页数	文献出版时占用期刊的页数。
	出版日期	出版日期	文献加工完成后到数据库的日期。
表名	表名	来源的 cnki 总库的对应数据库表名称	

读者推荐图片	读者推荐图片	由读者推荐的其他文献或期刊中的图片。
参考文献图片	参考文献图片	以文后参考书目或脚注的形式列出的文献题录中的图片。
引证文献图片	引证文献图片	引用作为参照文献的其他文献中的图片。
机构	机构	文献发表时作者所属机构的名称。
中文关键词	中文关键词	揭示单篇文献主题内容的一个或多个中文词。
中文摘要	中文摘要	揭示内容主题的一段中文文字。
英文篇名	英文篇名	文献标题的英语表述。
英文作者	英文作者	对文献实体或数据库记录负有责任的团体或个人名称的其他语种描述。
英文摘要	英文摘要	揭示内容主题的一段英文文字。
英文关键词	英文关键词	揭示单篇文献主题内容的一个或多个英文词或词组。
专题代码	专题代码	来源文献的专题代码
子栏目代码	子栏目代码	来源文献的子栏目代码
专题名称	专题名称	根据中图法分类法对文献进行的学科分类名称。
专题子栏目代码	专题子栏目代码	来源文献的专题子栏目代码
SYS_VSM	SYS_VSM	来源文献的语义特征向量
下载频次	下载频次	图片被下载的频率次数。
被引频次	被引频次	图片被引用的频率次数。
文献标识码	文献标识码	根据《中国学术期刊（光盘版）检索与评价数据规范》（CAJ-CD B/T1-1998）第4章关于文献标识码的规定，所确定的文献类型。
期刊标识码	期刊标识码	根据《中国学术期刊（光盘版）检索与评价数据规范》（CAJ-CD B/T1-1998）第15章关于期刊基本参数的规定格式编排的标识。
来源标识码	来源标识码	来源文献的文献标识码
作者代码	作者代码	来源文献的作者标识代码
机构代码	机构代码	来源文献的机构标识代码

10.2 数据收录标准

CNKI 图片知识库所有图片资源均收录自 CNKI 中国知识资源总库，其中主要来源于中国学术期刊网络出版总库、中国优秀硕士学位论文全文数据库、中国博士学位论文全文数据库、中国重要会议论文全文数据库、中国工具书网络出版总库、中国专利等。图片知识库是通过对上述资源的分析，挑选其中最具有价值的学术类图片整合成规范的知识库，现有图片超过 1500 万张，预计 2011 之后每年新增超过 100 万张。

其中图片收录的分布情况如下：

1. 《中国学术期刊网络出版总库》，是世界上最大的连续动态更新的中国学术期刊。截至 2010 年 10 月，收录全文文献总量 3000 多万篇；其中超过 63% 的文献中有规范的图片，

平均每篇含有图片的文献含有图片 4 左右；

2. 《中国优秀硕士学位论文全文数据库》，是国内内容最全、质量最高、出版周期最短、数据最规范、最实用的硕士学位论文全文数据库。截止至 2010 年 10 月，收录优秀硕士学位论文 107 多万篇；其中超过 80% 的优秀硕士论文中有图片，平均每篇论文中有图片 40 张左右；
3. 《中国博士学位论文全文数据库》，是国内内容最全、质量最高、出版周期最短、数据最规范、最实用的博士学位论文全文数据库。截止 2010 年 10 月，收录博士学位论文 13 万多篇。其中超过 80% 的博士论文中有图片，平均每篇的博士论文中有图片 50 张左右；
4. 《中国重要会议论文全文数据库》的文献是由国内外会议主办单位或论文汇编单位书面授权并推荐出版的重要会议论文。截至 2010 年 10 月，已收录累积文献总量 130 多万篇。其中超过 63% 的重要会议论文中有图片，平均每篇中有图片 4 张左右；
5. 《中国工具书网络出版总库》，简称《知网工具书库》，集成了近 200 家知名出版社的近 4000 余部工具书，其中医学图谱有 226 本，图录图鉴有 355 本；
6. 《中国专利全文数据库》包含国内各类的专利信息，准确地反映中国最新的专利发明。其中超过 63% 的专利中有图片，平均每个专利的文献中有 8 张左右图片。

10.3 各元数据质量标准

元数据标引质量指标：

表 10-2 元数据标引质量指标

元数据名称	质量指标	指标值
来源文献篇名	著录文字错误率	< 0.05%
来源文献作者	著录项错误率	< 0.05%
来源文献作者机构	著录项错误率	< 0.08%
来源文献刊名	著录项错误率	< 0.01%
年	著录项错误率	< 0.01%
期	著录项错误率	< 0.01%
页码	著录项错误率	< 0.05%
页数	著录项错误率	< 0.05%
来源文献中文摘要	著录文字错误率	< 0.1%
来源文献专题名称	著录项错误率	< 0.01%
来源文献文献标识码	著录项错误率	< 0.01%
来源文献期刊标识码	著录项错误率	< 0.01%
来源标识码	著录项错误率	< 0.01%

标引数据标引质量指标：

表 10-3 标引数据标引质量指标

元数据名称	质量指标	质量指标要求
图片中文标题	标引正确率	≥95
	标引一致度	≥95
图片说明	标引正确率	≥90%
	标引一致度	≥90%
图片关键词	标引正确率	≥90%
	标引一致度	≥90%
图片一级分类	标引正确率	≥90%
	标引一致度	≥90%
图片二级分类	标引正确率	≥90%
	标引一致度	≥90%
图片尺寸	标引正确率	≥99%
图片大小	标引正确率	≥99%

全文数据质量指标:

表 10-4 全文数据质量指标

项目名称	质量指标	质量指标要求
图像全文数据	文字和灰度图片扫描分辨率	≥300dpi
	彩色图片扫描分辨率	≥200dpi
	页面错误率	< 0.05%

10.4 内容分类标准与类别定义

10.4.1 图片内容分类体系构建说明

整合国家各类制图规范和标准,以及规范的出版物中对图片类型的定义与描述,制定出一套图片内容分类体系是本产品作为知识库产品的独有特点之一。

10.4.2 CNKI 图片知识库图片分类原则

1. 以 CNKI 中国知识资源总库文献中提取的图片为分类元数据;
2. 保证图片样例样图的清晰,内容完整,知识点表达明确;
3. 根据提取的图片图词与图片特征相结合的方式图片分类;以图片特征与关键词信息作为提取分类的图片载体;
4. 将图片和关键词信息按照特征分类、科学定义、关键词提取、样例样图查找的顺序进行分类;
5. 图词关键词的甄选过程保证图片关键词及其定义的规范性、权威性、科学性;
6. 图片特征与图词标题含义表达一致;图片特征的专业划分接近其图词所属学科;
7. 图片图词类别应尽量接近其所在基础学科;图词标引应与中国图书馆图书分类法相一致;

8. 此分类标准适用于 cnki 图片知识库为用户提供的图片检索和相似性图片检索的服务。

10.4.3 CNKI 图片知识库图片内容分类标准与类型定义

CT 图像：是 X 射线断层扫描技术简称。CT 图像通过一定数目的由黑到白不同灰度的像素，按矩阵排列所构成。

HPCE 图：高效毛细管电泳法。是以高压电场为驱动力，以毛细管为分离通道，依据样品中各组分之间淌度和分配行为上的差异而实现分离分析的液相分离方法。

Hill 图：根据希尔方程所得结合数据的图解表示方法。希尔方程是用于测定内缔合常数及每分子一定类型的蛋白质的结合部位的数目。它表明，肌肉在挛缩状态下，单位时间内从化学反应获得的机械能是常量；收缩力愈大，收缩速度愈小；反之亦然。

K 线图：又称蜡烛图、日本线、阴阳线、棒线等，常用说法是“K 线”，把每日或某一一周期的行情现况完全记录下来，所形成的连续的曲线即为 K 线图。

PAD 图：又称问题分析图，是一种算法描述工具。它是用二维数形结构的图表示程序的控制流，可以将这种图转换为程序代码表示一种程序逻辑。

TLC 图：又叫薄板层析，是色谱法中的一种，是快速分离和定性分析少量物质的一种很重要的实验技术，属于固—液吸附色谱。

TQM 图：即全面质量管理。为业务流程重建、全面质量管理和持续改进、创建因果、自上而下跨职能的流程图。

安装图：表示设备、构件等安装要求的图样。

饱和曲线：饱和曲线是根据击实曲线计算绘制的试样干密度与饱和含水率的关系推导的曲线。

饼图：排列在工作表的列或行中显示一个数据系列中各项的大小与各项总和的比例。饼图中的数据点显示为整个饼图的百分比。

波谱图：原子核或电子对频率的响应函数图形。

波形图：声波随时间变化呈现出不同的波幅。横轴表示时间，纵轴表示在单位时间内声波的波幅，这样形成的平面图即为波形图。

布局图：指从全局观点出发，对事物的全面规划和安排，从而得到的平面图。

布置图：平面布置图是工厂布置方案的一种简明图解形式，用以表示建筑物、构筑物、设施、设备等的相对平面位置。

彩色图像：能清楚的反应显影前的人或物的明确色彩的图象，即为彩色图像。

测试曲线：对在受控条件下运动的装备，进行功能和性能的检测，由此形成的曲线即为测试曲线。

层析图谱：是色层分析的简称。利用物质在固定相与流动相之间不同的分配比例，达到分离目的的技术。

超声心动图：用超声波检查心脏和大血管的解剖结构及活动状态的无创性技术。

等值线图：用一组相等数值点的连线表示连续分布且逐渐变化现象的数量特征的一种图型。等值线图又称等量线图。

滴定曲线：一种表示滴定过程中溶液性质变化的图示。

地图：按一定比例运用符号、颜色、文字注记等描绘显示地球表面的自然地理、行政区域、社会经济状况的图。

地形图：表示地表上的地物、地貌平面位置及基本的地理要素且高程用等高线表示的一种普通地图。

地质图：是表示地壳表层岩相、岩性、地层年代、地质构造、岩浆活动、矿产分布等的

地图的总称。

电镜图：用电子束为光源，显示标本超微结构的显微镜呈现的图像反射在平面上所形成的图像。

电路图：用导线将电源、开关（电键）、用电器、电流表、电压表等连接起来组成电路，再按照统一的符号将它们表示出来，这样绘制出的图形就叫做电路图。

电气图：用电气图形符号、带注释的围框或简化外形表示电气系统或设备中组成部分之间相互关系及其连接关系的一种图样。

电泳图谱：电泳分离后图形的记录，其形式可为电泳介质或载体本身或对其扫描等。

断面图：假想用剖切面剖开物体后，仅画出该剖切面与物体接触部分的正投影，所得的图形称为断面图，简称断面。

对比图：是把具有明显差异、矛盾和对立的事物双方安排在一起，进行对照比较的所形成平面图。

放电曲线：使带电的物体的电荷发生转移，正负电荷抵消使物体不显电性的一种变化曲线。

分布图：是表现一些现象空间分布位置与范围的图型。

分岔图：又称分叉图。用于可以将它比作树木长出枝条一样分支开来，以反映系统由一个过程向另一过程变化的图示；或表示动力学系统的参量值跨越临界值（分叉值）所导致稳定常状态定性变化的图示。

工程图：指由零件图或装配图构成的，用于机械设计、建筑工程施工等方面的二维图。

工况图：表示动力设备在一定条件下的工作状况或工艺条件，即各个参数之间的相互关系的图形。

工艺曲线：表示把各种原材料、半成品加工成为产品的方法和过程的曲线。

股价图：以特定顺序排列在工作表的列或行中用来显示股价的波动的数据。

股票走势图：股票走势图通常是叫做分时走势图或即时走势图，它是把股票市场的交易信息实时地用曲线在坐标图上加以显示的技术图形。坐标的横轴是开市的时间，纵轴的上半部分是股价或指数，下半部分显示的是成交量。分时走势图是股市现场交易的即时资料。分时走势图分为指数分时走势图和个股分时走势图。

关系曲线：事物之间的相互作用、相互影响的联系的状态。

管系图：表示管道系统中介质的流向，流经的设备，以及管件等连接，配置状况的图样。

光谱图：用摄谱仪获得的光谱记录。

函数曲线：表示每个输入值对应唯一输出值的一种对应关系的曲线。

荷载-位移曲线：显示荷载力值和位移数值之间比值的关系曲线。

横道图：又称甘特图 (Gantt chart)、条状图。是以图示的方式通过活动列表和时间刻度表示任何特定项目的活动顺序与持续时间。

户型图：户型图就是对房屋整体按一定比例压缩得到的一种模型图。它是对房屋整体形态的再现。

回归曲线：经固溶热处理的合金时效硬化后，在稍高于时效(低于固溶热处理)温度，短时间加热而引起性能复原的曲线。

回路图：回路图是用图形符号表示流体传动回路或部分回路功能的图。

肌电图：肌电图是记录神经和肌肉的电活动以判定神经和肌肉功能的一种检查方法。

级配曲线：将集料各级粒径颗粒的分配情况以曲线变化的表现的图示。

接线图：表示成套装置、设备或装置的连接关系的简图。

结构图：立体图案的一个组成部分，标示立体造型结构的组合形式和材料接合方法。

局部放大图：将图样中所表示的物体部分结构，用大于原图形的比例所绘出的图形。

竣工图：工程竣工验收后，真实反映建设工程项目施工结果的图样。

卡诺图：是一种平面方格图，又称为最小项方格图。卡诺图是逻辑函数的一种图形表示。一个逻辑函数的卡诺图就是将此函数的最小项表达式中的各最小项相应地填入一个方格图内，此方格图称为卡诺图。

控制图：控制图就是对生产过程的关键质量特性值进行测定、记录、评估并监测过程是否处于控制状态的一种图形方法。

框图：用线框、连线和字符表示系统中各组成部分的基本作用及相互关系的简图。

扩增图谱：指同一染色体区域的许多复制本通过 DNA 复制，结果产生小卫星样染色体或染色体 DNA 区域出现多套基因复制本，从而形成的一种图谱。

雷达图：又称为戴布拉图、蜘蛛网图。即将一个公司的各项财务分析所得的数字或比率，就其比较重要的项目集中划在一个圆形的固表上，来表现一个公司各项财务比率的情况。

立面图：建筑物、构筑物等在直立投影上所得的图形。

立体图：以立体几何体积的大小或立体形的长、宽、高度来表示统计指标数值大小的图形。

淋洗曲线：淋洗曲线是以淋出的溶液体积为横坐标，目标的组分的量为纵坐标画的曲线。可以看出目标的组分在淋出时随淋洗液体积的变化情况。

零件图：是表示单个零件形状、大小和特征的图样，也是在制造和检验机器零件时所用的图样，又称零件工作图。

流程图：用规定的符号和简单的文字说明来描述数据处理步骤的图式。它借助图线和箭号表示其处理步骤的流向或顺序。

逻辑图：用二进制逻辑单元图形符号所绘制的电路简图。逻辑图由许多逻辑图形符号构成。

毛坯图：零件制造过程中，为铸造，锻造等非切削加工方法制作坯料时提供详细资料的图样。

酶切图谱：描述限制性内切酶的酶切点的位置和距离信息的图谱。

面积图：排列在工作表的列或行中强调数量随时间而变化的程度，用于引起人们对总值趋势的注意。

模块图：将系统的设备和设施按照一定的结构或形式，构成具体一致关系的图形。模块图，又称构件图。

模态分析图：是结构的固有振动特性，每一个模态具有特定的固有频率、阻尼比和模态振型。这些模态参数可以由计算或试验分析取得，这样一个计算或试验分析过程称为模态分析。

模型图：它是所研究的系统、过程、事物或概念的一种表达形式，也可指根据实验、图样放大或缩小而制作的样品，一般用于展览或实验或铸造机器零件等用的模子。

挠度曲线：结构构件的轴线或中面由于弯曲引起垂直于轴线或中面方向的线发生位移而形成的曲线。

脑电图：是将脑部自发性生物电活动通过电子放大技术描记于纸上而得到的图形。

能谱图：海浪可视作由无限多个振幅不同、频率不同、方向不同、位相杂乱的组成波组成，这些组成波便构成海浪谱。它用于描述海浪内部能量相对于频率和方向的分布，故又名“能量谱”、“功率谱”和“方向谱”。

鸟瞰图：用高视点透视法绘制的地图。使视线与水平线有一俯角，读者似从高处俯视制图区。

配置图：又称部署图，是用来显示系统中软件和硬件的物理架构。使用部署图可以显示运行时系统的结构，同时还传达构成应用程序的硬件和软件元素的配置和部署方式。

频谱图：以横轴纵轴的波纹方式，记录画出各种信号频率的图形资料。

剖面图：假想以垂直于地表的截面切割地面以反映地面起伏曲线或内部构成的图形。

气象图：气象图也叫天气图。根据不同要求和目的而有多种类别。通常专指根据同一时刻测得的各地天气实况，用数值或符号按一定格式填在空白地图上，反映广大地区天气实况和形势的图。

切面图：指将一个器官沿一个方向切开，以显示其结构的图片称为切面图。

趋势图：也可称为统计图或统计图表。是以统计图的呈现方式来表现某物或某信息数据的发展趋势的图形。

曲面图：排列在工作表的列或行中的显示两组数据之间的最佳组合情况。

全息图：以激光为光源，用全景照相机将被摄体记录在高分辨率的全息胶片上构成的图。全息图是一种三维立体图像。这是一种像胶片相片一样的储存图像信息的媒介。

扫描图：通过电子束、无线电波等的左右移动在屏幕上显示出来的画面或图形。

色谱图：色谱柱流出物通过检测器系统产生的响应信号对时间或载体流出体积的曲线图。

设计图：在工程项目或产品进行构形和计算过程中所绘制的图样。

声像图：以光点亮暗显示超声波扫描机体被查部位器官和组织结构的二维断层图像，如B超。

失重曲线：物体在零重力的情况下，在引力场中自由运动时有质量而不表现重量的一种曲线状态。

施工图：是表示建筑物或设备、管、线的各部分的布置、作法、构造、相互间的关系、施工及安装要求、质量要求的详细图纸和说明。

时程曲线：通过积分运算求得在地面加速度随时间变化期间结构的内力和变形状态随时间变化的全过程。

实测曲线：是指在未开展同比例尺工作的范围，按相应技术精度要求开展的填图工作。

示功图：所包围的面积可表示为机器所作的功或所消耗的功，故称为示功图，它可用示功器测录。

示意图：表达机械的工作原理，如表达机械传动原理的机械运动简图、表达液体或气体输送线路的管道示意图等。

视图：是按正投影法即机件向投影面投影得到的图形。按投影方向和相应投影面的位置不同，视图分为主视图、后视图、俯视图和左视图等。

数据流图：简称DFD，它从数据传递和加工角度，以图形方式来表达系统的逻辑功能、数据在系统内部的逻辑流向和逻辑变换过程，是结构化系统分析方法的主要表达工具及用于表示软件模型的一种图示方法。

衰减曲线：以规定的形式，用两点功率之比，或与功率有关的量之比值来表示的功率的衰减，从而形成衰减曲线。

顺序图：是将纵向时间轴代表时间，横向轴代表各独立对象的类元角色，表示角色之间具有一定的序列关系的二维图。

思维导图：又称心智图，思维导图是表达发射性思维的有效图形思维工具。

梯形图：梯形图常被称为电路或程序，梯形图的设计称为编程。梯形图与电器控制系统的电路图很相似，特别适用于开关量逻辑控制。

条形图：亦称纵条图、柱形图、带形图、条状图。以宽度相等的条形的高低或长短表现统计数字大小和变动的一种统计图式。现在默认条形图和条状图是一回事。

统计图：是利用点、线、面、体等绘制成几何图形，以表示各种数量间的关系及其变动情况的工具。

透视图：在人与建筑物之间设立一个透明的铅垂面K作为投影面，人的视线（投射线）

透过投影面而与投影面相交所得的图形，称为透视图，或称为透视投影，简称透视。

网络图：是一种图解模型，形状如同网络，故称为网络图。网络图是用箭线和节点将某项工作的流程表示出来的图形。

位移曲线：在平面直角坐标系中，以纵轴为位移轴，横轴为时间轴，画出位移和时间的关系图线，这种图线称为位移-时间曲线，简称为位移曲线。

卫星图像：是各种人造地球卫星在运行过程中，通过照相机、电视摄像机、多光谱扫描仪等设备，对地面地物进行摄影或扫描所获得的图像资料，有时也称卫星像片。

温谱图：在程序升温或降温下，测定物质的物理参数与温度之间的关系的图谱。

熵焓图：以气温 (t) 为横坐标，位温对数($\ln\theta$)或熵(S)为纵坐标的热力图。

误差曲线：表示计量器具误差与被测量之间、或与对比误差有影响的任何其它量之间函数关系的曲线。

洗脱图：以柱层析洗脱物的某种性质如吸光率或放射性作为洗脱物分段数，或收集洗脱物累积体积的函数所作的图。

线型图：将点、横线和空格等按一定规律重复出现形成的图形即为线型图。

相关图：又称散布图。把相关的两个与质量有关的因素的数据，用点子描绘在坐标纸上，以观察两个因素之间关系的图解。

详图：表明生产过程中所需要的细部构造、尺寸及用料等全部资料的详细图样。

响应曲线：根据遥感器对波谱的相对响应与波长的关系在直角坐标系中描绘出的曲线。

效果图：可以理解为对设计者的设计意图和构思进行形象化再现的形式。

心电图：亦称“体表心电图”或“常规心电图”。将心脏激动过程的微弱电流，经人体组织传至体表，通过心电图记录仪描记下来，形成动态曲线，即为心电图。

形貌图：物体表面或者内部形态发生变形的形态图。

性能曲线：通过测试工具模拟多种正常、峰值以及异常负载条件来对系统的各项性能指标进行测试而得到的曲线。

序列图：指描述数据序列的图。

学习曲线：学习曲线有时又称练习曲线，人们为了知道学习进程中的现象和进步的快慢的详情，作为以后努力的指针，应用统计图的方法作一条线，把它表示出来。

血流图：是在检查部位引入高频率、低电压的弱电流后，观察血流变化所引起阻抗变化的图形。

循环图：显示循环的图表。

压力曲线：指将受测者置于某一特定的极端市况下，然后测试该受测者在这些关键变量突变的压力下的表现状况，看是否能经受得起这种情况的突变，从而得到压力曲线。

衍射图谱：波遇到障碍物或小孔后通过散射继续传播获得的图谱。

样图：根据规范、图式或者编图大纲等，将典型地区制成的编绘样图。

遥感图像：或称遥感像片，是各种传感器所获信息的产物，是遥感探测目标的信息载体。

因果图：又称鱼骨图、石川图、因果图、Ishikawa。问题的特性总是受到一些因素的影响，我们通过找出这些因素，并将它们与特性连在一起，按相互关联性整理而成的层次分明、条理清楚，并标出重要因素的图形就叫因果图。

影响图：项目管理中用于识别风险，显示因果影响，按时间顺序排列的事件以及变量与结果之间的其他关系的图解表示法。

语谱图：语谱图即语音频谱图，一般是通过处理接收的时域信号得到频谱图，得到频谱分析视图，如果这种频谱分析视图是针对语音数据的话，叫语谱图。

原理图：用规定的符号、代号和简化画法绘制，表示产品组成部分的结构、动作等原理的图样。

圆环图：排列在工作表的列或行中显示各个部分与整体之间的关系。区别于饼图，它可以包含多个数据系列。

云图：地球上不同尺度的云层覆盖和充满变化的地表面特征的图像。

造影图：通过注射或口服某些 X 射线不能透过的药物，使某些器官在 X 射线下显示出来，以便对疾病进行检查。

照片：用照相术拍得的画面、影像、肖像，亦称“相片”。用照相术拍得的画面、影像、肖像。亦称“相片”。

折线图：排列在工作表的列或行中显示随时间而变化的连续数据，显示在相等时间间隔下数据的趋势。

指纹图谱：又称肽链图谱。是一种结合电泳和色谱法分析肽链氨基酸组成的方法。这种方法对于研究人和动物 Ig 以及研究人的血红蛋白变种和病毒变异都很有用。

质粒图谱：用以记录细菌细胞内一种环状双链 DNA 分子的自我复制的过程的图形。它能稳定地独立存在于染色体外，并传递到子代，一般不整合到宿主染色体上。

滞回曲线：它反映结构、构件或岩土试件在反复受力过程中的变形特征、刚度退化及能量消耗，是确定恢复力模型和进行非线性地震反应分析的依据。又称恢复力曲线。

轴测图：用平行投影法将空间形体和确定其位置的空间直角坐标系投影到投影面上得到的图形。

柱状图：也称条图、长条图、条状图。是一种以长方形的长度为变量的表达图形的统计报告图，由一系列高度不等的纵向条纹表示数据分布的情况，用来比较两个或以上的价值，只有一个变量，通常利用于较小的数据集分析。饼图也是柱形图的一类。

装配图：表示产品及其组成部分的连接、装配关系的图样。

装修图：装修图是装修设计图片的简称。装修设计图片指用点、线、面、符号、文字和数字等描绘空间建筑物几何特征、形态、位置及大小的一种形式。装修图片可以分为建筑装修施工图、装修效果图和装修实景图等。

装置图：指机器、仪器或其他设备等具有某种独立功用的物件之间关系的图。

11 专家库数据规范

11.1 专家库元数据

11.1.1 元数据及其定义

《专家库》的元数据及其定义见下表：

表 11-1 专家库元数据及其定义

元素名称	名称	标签	定义
Title	专家姓名	专家姓名	单条记录的主体对象名称。
	其他语种专家姓名	其他语种专家姓名	单条记录主体对象名称的其他语种描述。包括其他语言作者姓名等。
Subject	科研领域	科研领域	揭示专家研究内容的一个或多个 CNKI168 专题
	学科	学科	揭示专家研究内容的一个或多个国家国务院学位办学科分类

元素名称	名称	标签	定义
	专家关键词	关键词	揭示专家研究内容的一个或多个规范词
Description	当前单位	当前单位	专家信息发表时作者所属机构的名称
	当前职称	当前职称	专家当前的专业技术职称
	基本信息	基本信息	专家个人简介信息，包括专家的性别、出生日期、出生地，学历、职称、职务等信息。
	研究方向	研究方向	揭示专家专业、领域简短的描述性的短语或短句
	履历	履历	专家的个人经历的简要说明，一般描述专家在什么时间干了什么事情。 个人经历包括：教育背景，工作履历，学术兼职，社会兼职
	荣誉	荣誉	专家所获奖项信息或者荣誉称号、头衔。一般指政府、社团、所属单位、其他组织或机构给予专家的一种积极的正式的评价或称号。
	联系方式	联系方式	专家的联系信息，一般包括不限于地址，邮编，电话，手机，mail，,MSN，QQ，,微薄帐号，个人主页，博客
	论文	论文	专家公开发表的学术文献
	著作	著作	专家公开发表的著作，一般包括不限于图书、教材等。
	报告	报告	专家参与国际、国内重大学术会议的报告
	成果	成果	专家的工作、事业、科研方面的成绩或成效描述。此处的成果是指不包括论文，著作，报告以外的专家的各项成果，且有具体的成果条目或成果描述语言。一般包括不限于指标准，专利，学术创新点，科学研究成果，网络文集，百科辞典条目，年鉴条目，专家带过的学生，专家教过的课程，课件，艺术创作作品，参与创作，编排的电影、电视音乐作品等
	科研项目	科研项目	专家的科研经历，科学研究信息及科研项目信息
	所获基金资助	所获基金资助	专家完成科研项目过程中所受资助的基金名称及资助说明
	社会活动	社会活动	专家参与社会活动的经历。一般包括不限于：学术论坛、标准制定会、政策法规论证、学术报告、演讲、人大会议、法庭辩论、价格听证会等。
	成果荣誉综述	成果荣誉综述	专家的个人论文、著作、报告、成果、所获荣誉、科研项目、社会活动等信息的整体的、综合性的介绍或简要概述。
	个人综合报道	个人综合报道	公开发表的专家的人物专访 公开发表的同行评议或评论，可以对专家个人的评价也可以是对专家某个观点的评论
照片	照片	专家的影像或肖像，一般是为专家单人正面照。	
Identifier	专家编码	专家编码	由专家姓名和专家机构确定的专家的唯一标识符
	机构编码	机构编码	专家机构对应的标准一级机构编码
Relation	导师编码	导师编码	专家的博士导师，硕士导师对应的专家编码

元素名称	名称	标签	定义
	学生编码	学生编码	专家所带的博士生，硕士生对应的专家编码

11.1.2 元数据的规范控制

对表 11-1 中的部分元数据使用适当的规范文档，以便对其检索功能进行规范控制。可以根据具体情况选择前控方式或后控方式。

1. 对于专家姓名，使用作者规范文档进行规范控制；
2. 对于专家关键词，使用主题词表进行规范控制；
3. 对于机构，使用机构规范文档进行规范控制。

11.1.2.1 元数据标引

对每位专家给出一些可以作为检索标识的元数据。

11.1.2.2 标引元数据项

标引元数据项包括：科研领域、学科，专家关键词、当前单位、当前职称、研究领域、基本信息、履历、论文、著作、报告、成果、科研项目、社会活动、所获荣誉、作品成果荣誉综述、个人综合报道、联系信息。

11.1.2.3 标引要求

1. 准确、全面地进行标引；
2. 专家科研领域来自 CNKI168 专题体系，根据 CNKI 收录的专家文献所属 168 专题统计计算得出；
3. 学科来自中国国家国务院学位办学科分类标准，根据 CNKI 收录的专家文献所属学科计算得出。

11.1.2.4 质量指标

对于反映专家的标引元数据，有两个质量指标：标引正确率和标引一致度。

11.1.2.5 质量指标要求

各标引元数据的质量指标要求见下表：

表 11-2 元数据标引质量指标

元数据名称	质量指标	质量指标要求
科研领域	标引正确率	≥90%

元数据名称	质量指标	质量指标要求
	标引一致度	≥90%
学科	标引正确率	≥90%
	标引一致度	≥90%
当前单位	标引正确率	≥90%
	标引一致度	≥90%
基本信息	标引正确率	≥90%
	标引一致度	≥90%
履历	标引正确率	≥90%
	标引一致度	≥90%
论文	标引正确率	≥90%
	标引一致度	≥90%
著作	标引正确率	≥90%
	标引一致度	≥90%
报告	标引正确率	≥90%
	标引一致度	≥90%
成果	标引正确率	≥90%
	标引一致度	≥90%
科研项目	标引正确率	≥90%
	标引一致度	≥90%
社会活动	标引正确率	≥90%
	标引一致度	≥90%
所获荣誉	标引正确率	≥90%
	标引一致度	≥90%
作品成果荣誉综述	标引正确率	≥90%
	标引一致度	≥90%
个人综合报道	标引正确率	≥90%
联系信息	标引正确率	≥90%

11.2 专家知网节的内容

11.2.1 节点专家

11.2.1.1 专家姓名

链接目标：分别链接网上同名专家。

功能：发现国内外同名专家。

11.2.1.2 专家所在机构

链接目标：链接同领域同机构的专家。

功能：全面反映同机构中相同领域专家情况。利用专家名望、机构品牌的影响力传播。

11.2.1.3 专家关注点（关键词）

链接目标：推荐可能感兴趣的专家，推荐同领域同研究兴趣的专家。

功能：发现可能存在的新知识；发现同研究兴趣的其他专家的研究进展，发现国内外重复、竞争性研究项目。发现同领域相同研究兴趣的专家。

11.2.1.4 专家科研领域及排序

链接目标：介绍专家科研领域及专家在默认领域的排序情况。

功能：显示专家在本领域的学术影响力及地位，包括专家的发文、被引频次、H 指数、G 指数在本领域的排序。

11.2.1.5 专家介绍

链接目标：介绍专家基本信息。

功能：显示专家基本信息，包括专家基本信息，专家履历，专家的研究方向，专家荣誉，专家联系方式，使用户能快速了解专家。

11.2.1.6 科研产出

链接目标：链接专家科研产出，包括论文，著作，专利，其它成果综述。

功能：全面展示专家的所有学术成果。反映专家的科研生产力。

11.2.1.7 影响力统计分析

链接目标：链接专家的学术影响力，包括发表文献统计，被引频次统计分析，所在领域年度排序，所受基金资助资助情况统计。

功能：全面展示专家的学术影响力。从多个反映专家在本领域的学术。

11.2.1.8 学术关系网络

链接目标：链接专家的学术关系网络。

功能：展示专家学术关系，包括专家合作关系，专家引用关系，专家师生关系。

11.2.1.9 他人眼中的专家

链接目标：链接对专家的综合评价。

功能：展示对专家的综合评价，包括专家情况的综述，专家个人专访以及专家参与的社会活动。

11.2.1.10 所在领域的研究现状

链接目标：链接本领域的学术影响力排序前 300 的专家。

功能：全面反映本领域的学术影响力较高的专家，包括发文排序、被引排序、H 指数排序及 G 指数排序前 300 的专家。

11.2.2 节点专家的相关专家连接

11.2.2.1 同领域同单位专家

与节点专家的关系：与本专家（节点专家）同单位相同科研领域的专家。

功能：了解本专家的同单位与其相同科研领域的专家。

11.2.2.2 同领域博导

与节点专家的关系：与本专家相同科研领域的博士生导师。

功能：了解本领域的博士生导师的情况。

11.2.2.3 同领域硕导

与节点专家的关系：与本专家相同科研领域的硕士生导师。

功能：了解本领域的硕士生导师的情况。

11.2.2.4 可能感兴趣的专家

与节点专家的关系：用户感兴趣的专家。

功能：根据用户使用专家库模式推荐用户感兴趣的专家。

11.2.2.5 同领域同研究兴趣的专家

与节点专家的关系：与本专家相同领域相同研究兴趣的专家。

功能：揭示本领域其它的专家。

11.2.2.6 同领域同基金资助的专家

与节点专家的关系：与本专家相同科研领域且得到过相同基金资助的专家。

功能：揭示本领域其它的专家。

11.3 专家收录的完整性指标

11.3.1 专家库完整性质量指标

11.3.1.1 名家收集完整率

专家库实际收集的名家数与国内实际名家数之比。

名家收集完整率不低于 98%。

11.3.1.2 正高级职称专家收集完整率

专家库中实际收集的正高职专家与中国专家学者库中正高级职称的专家之比, 收集完整率不低于 70%。

11.3.1.3 各科研领域专家收集完整率

完整率不低于 60%。

11.3.2 专家信息收集完整率

专家信息收集完整率不低于 60%。

12 作者规范

作者信息包含：姓名、性别、出生日期和简历等信息。

12.1 作者姓名

作者姓名应包含有该作者的正式名称、现用名、曾用名、笔名、汉语拼音拼写等一切可以标识该作者的名称。

12.1.1 作者姓名的选择

对于曾用名录入，要求记录使用该名称的时间以及该作者所有曾经使用过的名称。

12.1.2 作者名称的书写

汉语拼音的书写要求是姓名的全拼，姓的首字母和名的首字母必须大写，姓与名中间用“-”相连。

12.2 性别

遵循 GB/T 2261.1-2003《个人基本信息分类与代码》第 1 部分：人的性别代码的规定。

12.2.1 简历

简历可提供职务信息、职称信息、学历等信息。

13 引文数据库规范

13.1 元数据

13.1.1 元数据及其定义

《中国引文数据库》的元数据及其定义如下表 13-1 所示：

表 13-1 《中国引文数据库》元数据及其定义

元素名称	名称	标签	定义
题名	中文题名	题名	被引文献的题名。包括正题名、副题名、并列题名、副书名、分辑名等。
	其他语种题名	其他语种题名	被引文献题名的其他语种描述。包括其他语言正题名、副题名、并列题名等。
作者	作者中文名	作者中文名	对被引文献实体或数据库记录负有责任的团体名称或个人名称。
	其他语种作者姓名	其他语种作者姓名	对被引文献实体或数据库记录负有责任的团体名称或个人名称的其他语种描述。
	第一责任人	第一作者	被引文献的第一作者。
	作者机构	作者机构	被引文献发表时作者所属机构的名称。

元素名称	名称	标签	定义
其他 责任 者	导师	导师	对该被引的博士（硕士）论文进行指导高等学校或研究机关中的教师或科研人员。
	第一导师	第一导师	该被引博士（硕士）论文的第一导师。
	导师单位	导师单位	指导该篇被引的博士（硕士）论文时导师所属机构名称。
主题	中文关键词	关键词	揭示单篇被引文献主题内容的一个或多个中文词。
	英文关键词	英文关键词	揭示单篇被引文献主题内容的一个或多个英文词或词组。
	中国图书馆分类法分类号	分类号	按照中国图书馆分类法体系，根据被引文献内容所涉及的主题指定的类别号。
	中国图书馆分类法名称	分类名称	按照中国图书馆分类法体系，根据被引文献内容所涉及的主题制定的类别名称。
	学科专业名称	学科专业名称	(?)
	类型标识	自社科	用以标识来源期刊类型，分为两种类型：自科技和自社科。
描述	中文摘要	摘要	揭示被引文献内容主题的一段中文文字。
	英文摘要	英文摘要	揭示被引文献内容主题的一段英文文字。
	基金	基金	在完成被引文献或被引文献所属项目过程中所受资助的基金名称及资助说明。
	版本说明	版本说明	被引文献类型为图书的版本说明。
日期	网络出版日期	网络出版日期	被引文献加工完成后出版到数据库中的日期。
出版 者	出版单位	出版单位	使印刷版文献成为可取得或利用状态的责任者。
	出版地	出版地	印刷版出版者注册所在地，根据中国行政区划划分。
	学会名称	会议名称	被引会议论文所属会议主办方名称。
	会议地点	会议地点	被引会议论文所属会议召开地点，根据中国行政区划划分。
标识 符	文献标识码	文献标识码	根据《中国学术期刊(光盘版)检索与评价数据规范》(CAJ-CD B/T 1-1998)第4章关于文献标识码的规定，所确定的文献类型。
	数字对象统一标识符	DOI	对数字对象统一定义的唯一标识符。
	文献统一编号	文件名	根据《中国学术期刊(光盘版)检索与评价数据规范》(CAJ-CD B/T 1-1998)中第5章关于文献编号的规定，为文献指定的、全世界范围内唯一的文献编号。
	期刊标识	期刊范围	用以标识期刊是否属于“核心期刊”。
	ISBN	ISBN	被引的所有正规出版的普通图书类型文献的国际标准书号。
	专利号	专利号	被引专利类型文献的统一标识码。
	标准号	标准号	被引标准类型文献的统一标识码。
	引证标识	引证标识	用以区分文献的参考和引证情况，分三种类型：仅参考、仅引证、两者皆有。
来源	中文刊名	来源文献	中文版期刊的正式名称，或外文版期刊的正式中文翻译名称。对应于“来源名称”。
	英文刊名	英文刊名	英文版期刊的正式名称，或非英文版期刊的正式英文翻译名称。

元素名称	名称	标签	定义
	拼音刊名	拼音刊名	(?)
	出版年卷期号	出版年卷期号	期刊出版的年份、卷号及期号。
	期刊基本参数	期刊基本参数	根据《中国学术期刊(光盘版)检索与评价数据规范》(CAJ-CD B/T 1-1998)中第15章关于期刊基本参数的规定,包含该期刊某些参数并按规定格式编排的标识。
	会议名称	会议名称	被引会议论文所属会议名称。
	学位授予单位	学位授予单位	该被引博士(硕士)论文责任者的博士(硕士)学位授予机构。
	国际标准连续出版物号	ISSN	根据国际标准组织1975年制订之ISO-3297的规定,由设于法国巴黎的国际期刊资料系统中心(International Serial Data System - ISDS International Centre)所赋予,申请登记的具有唯一性、可识别性、国际通行性的统一编号。
	中国出版物编号	CN	中国出版物标准编号。
	电子版国际标准连续出版物号	ISSN	根据国际标准组织1975年制订之ISO-3297的规定,由设于法国巴黎的国际期刊资料系统中心(International Serial Data System - ISDS International Centre)所赋予申请登记的具有唯一性、可识别性、国际性的统一编号。
	中国连续电子出版物编号	CN	中国电子出版物标准编号。
	电子版年期号	电子版年期号	以光盘形式发布时,电子文件所在光盘的编号。
语种	语种	语种	文献所使用的主要语言。
关联	参考文献	参考文献	以文后参考书目或脚注的形式列出的文献题录。
	引证文献	引证文献	引用作为参照的文献的其他文献(此项为动态元数据元素)。
	共引文献	共引文献	与本文引用相同文献的文献。(此项为动态元数据元素)
	同被引文献	同被引文献	与本文同时被其它文献引用的文献(此项为动态元数据元素)。
	二级参考文献	二级参考文献	本文参考文献的参考文献。(此项为动态元数据元素)
	二级引证文献	二级引证文献	本文引证文献的引证文献。(此项为动态元数据元素)

13.1.2 元数据的规范控制

对表 13-1 中的部分元数据使用适当的规范文档,以便对其检索功能进行规范控制。可以根据具体情况选择前控方式或后控方式。

- a) 对于期刊类型引文当中的来源,即期刊名称,使用期刊名称规范文档进行规范控制;
- b) 对于文后参考,使用参考文献规范文档进行规范控制。

13.2 元数据著录规则及其质量要求

元数据包括中文题名、其他语种题名、作者中文名、其他语种作者姓名、第一责任人、

作者机构、导师、第一导师、导师单位、中文关键词、英文关键词、中国图书馆分类法分类号、中国图书馆分类法名称、学科专业名称、类型标识、中文摘要、英文摘要、基金、版本说明、网络出版日期、出版单位、出版地、学会名称、会议地点、文献标识码、数字对象统一标识符、文献统一编号、期刊标识、ISBN、专利号、标准号、引证标识、中文刊名、英文刊名、拼音刊名、出版年卷期号、期刊基本参数、会议名称、学位授予单位、国际标准连续出版物号、中国出版物编号、电子版国际标准连续出版物号、中国连续电子出版物编号、电子版年期号、语种、参考文献、引证文献、共引文献、同被引文献、二级参考文献、二级引证文献。

引文库的元数据质量，要符合各来源全文库元数据的著录质量指标要求。

13.3 引文数据库规范文档要求

13.3.1 参考文献规范

参考文献规范参考了国家新闻出版总署颁发的《中国学术期刊（光盘版）检索与评价数据规范》及国家质量监督检验检疫总局和国家标准化管理委员会共同发布的《文后参考文献著录规则》。

- a) 参考文献所指文献若已在 CNKI 收录的全文或题录数据库中，应在客观公正地遵照作者书写参考文献格式的基础上，与全文或题录数据库的文献来源信息对比，规范责任者、题名、来源、时间、版本、出版项、页码等信息；
- b) 参考文献所指文献未在 CNKI 收录的全文或题录数据库中，应主要规范参考文献信息中的来源项。

13.3.2 期刊名称规范

规范参考文献信息中的来源信息，主要规范对象是期刊类型参考文献的期刊名称。期刊名称的规范应遵循国家新闻出版总署颁发的新出报刊[1999]1115号《关于规范期刊刊名标识的通知》的规定。

当期刊更名时，CNKI 数据库采用最新的期刊名称，并保证新旧刊名对应一致；当期刊有简称、俗称时，CNKI 数据库一律采用该期刊的标准名称，并保证标准名称与简称、俗称等对应一致；

对于 CNKI 各库中未收录的期刊，采用最新的期刊名称表，并对各种缩写、简称、刊名变更情况进行标准化处理，保证期刊统计数据的完整性和准确性。

13.4 资源收录原则与规范

引文库的资源来源于《中国学术期刊网络出版总库》下学术期刊的参考文献，文献类型包括：期刊类型、图书类型、学位论文类型、会议论文类型、专利类型、标准类型、报纸类型、年鉴类型等。从资源来源角度分析，引文库来源期刊的参考文献分为两部分：CNKI 各源数据库产品已收录全文或题录资源的参考文献，CNKI 各源数据库产品未收录资源的参考文献。

13.4.1 已收录全文或题录资源的参考文献部分的收录原则

CNKI 已收录全文或题录资源的参考文献被引文库收录，提供检索和统计分析功能。

13.4.2 未收录资源的参考文献部分的收录原则

CNKI 未收录资源的参考文献部分的资源，根据资源类型作不同方式的处理。

未收录资源的期刊类型参考文献

规范参考文献信息中的期刊名称；在源文献检索和引文检索中均对期刊刊名提供检索，检索结果客观显示刊名信息（以原文作者在参考文献列表中所写的期刊刊名为准）；可对规范后的该期刊进行统计分析。参考文献数据的其他部分如作者信息、题名信息等不做规范，也不提供检索和统计。

未收录资源的其他类型参考文献

引文库不予收录，在源文献检索和引文检索中均不提供检索和统计。