



中华人民共和国国家标准

GB/T 13190.1—2015

代替 GB/T 13190—1991, GB/T 15417—1994

信息与文献 叙词表及与其他词表的互操作 第 1 部分：用于信息检索的叙词表

Information and documentation—Thesauri and interoperability
with other vocabularies—Part 1: Thesauri for information retrieval

(ISO 25964-1:2011, MOD)

2015-05-15 发布

2015-12-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

中 华 人 民 共 和 国
国 家 标 准
信息与文献 叙词表及与其他词表的互操作
第 1 部分：用于信息检索的叙词表

GB/T 13190.1—2015

*

中国标准出版社出版发行
北京市朝阳区和平里西街甲 2 号(100029)
北京市西城区三里河北街 16 号(100045)

网址：www.gb168.cn

服务热线：400-168-0010

010-68522006

2015 年 6 月第一版

*

书号：155066·1-51535

版权专有 侵权必究

目 次

前言	V
引言	Ⅶ
1 范围	1
2 规范性引用文件	1
3 术语和定义	2
4 叙词表概述和目标	11
4.1 总体目标	11
4.2 词汇控制及目的	11
4.3 聚合关系和组合关系	12
4.4 聚合关系的类型	12
4.5 符号、简称和其他约定	13
5 叙词表中的概念及其范围	15
5.1 概念基础	15
5.2 范围注释	16
5.3 互见范围注释	17
6 叙词表语词	17
6.1 语词形式	17
6.2 叙词表语词的明晰与消歧	17
6.3 语词的语法形式	19
6.4 大小写、标点符号和特殊字符	21
6.5 单数或复数形式	23
6.6 优选词的选择	24
7 复杂概念	31
7.1 概述	31
7.2 复合语词的特点	32
7.3 确定是否采用复杂概念	33
7.4 怎样分解复杂概念	36
7.5 保留优选词的组成概念	36
7.6 处理复杂概念的一致性	37
7.7 多单词语词中的单词顺序	37
8 单语种叙词表中的等同关系	38
8.1 概述	38
8.2 同义词	38
8.3 准同义词	39
8.4 包含在上位概念中的专指词	40
8.5 用组合语词表示复杂概念	40

9	跨语言等同	41
9.1	概述	41
9.2	等同程度	41
9.3	典型问题及解决办法	42
9.4	优选词之间的跨语言等同表达方式	46
9.5	非优选词之间的跨语言等同	46
10	概念之间的关系	47
10.1	概述	47
10.2	等级关系	47
10.3	相关关系	51
10.4	自定义关系	55
11	分面分析	55
12	显示与布局	57
12.1	概述	57
12.2	显示方式	58
12.3	多语种叙词表显示与布局	70
12.4	语言和字符编码问题	77
13	叙词表的构建与维护管理	78
13.1	叙词表规划	78
13.2	编制工作初期阶段	80
13.3	构建叙词表	81
13.4	介绍叙词表	82
13.5	发布	83
13.6	更新	84
14	叙词表管理软件的指导原则	87
14.1	概述	87
14.2	规模和字符限制	87
14.3	语词间和概念间的关系	87
14.4	语词和概念的注释	88
14.5	代码和符号	88
14.6	节点标签	88
14.7	语种地位	88
14.8	数据导入与导出	88
14.9	编辑性导航和支持	89
14.10	编辑性保护措施	90
14.11	内部管理工具	90
15	数据模型	90
15.1	总则	90
15.2	模型说明	92
15.3	列表呈现	96
16	叙词表与应用程序的集成	101
16.1	简介	101

16.2	叙词表互操作的需求	101
16.3	与标引和检索应用程序集成	101
17	交换格式	103
18	协议	103
18.1	概述	103
18.2	目的和用途	103
18.3	应用环境和架构	104
18.4	叙词表专门协议	104
18.5	叙词表使用的通用 Web 数据库协议	104
附录 A (资料性附录)	国外已出版叙词表显示样例	105
附录 B (资料性附录)	国内已出版叙词表显示样例	120
附录 C (资料性附录)	用于数据交换的 XML 框架	132
参考文献	133

前 言

GB/T 13190《信息与文献 叙词表及与其他词表的互操作》分为以下两个部分：

——第1部分：用于信息检索的叙词表；

——第2部分：与其他词表的互操作。

本部分为 GB/T 13190 的第1部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

本部分代替 GB/T 13190—1991《汉语叙词表编制规则》、GB/T 15417—1994《文献 多语种叙词表编制规则》。与 GB/T 13190—1991、GB/T 15417—1994 相比，主要技术内容变化如下：

——增加了基于数字技术的词表编制规则(见 14.8、14.9、14.10)；

——增加了网络环境下的词表应用规则(见第 15 章、第 16 章、第 17 章、第 18 章)。

本部分使用重新起草法修改采用 ISO 25964-1:2011《信息与文献 叙词表及与其他词表的互操作 第1部分：用于信息检索的叙词表》。

本部分与 ISO 25964-1:2011 的主要技术差异如下：

——增加了“规范性引用文件”章，相应地将“术语和定义”调整为第 3 章，第 3 章的“符号、简称和其他约定”变为第 4 章的 4.5。

——为满足汉语叙词表编制的需要，在各章节增加汉语语词的举例，或以“说明(或注)”的形式说明汉语语词的特点。例如：

3.36 多单词语词

注：在汉语中，包括多个单字组成的语词，也包括多个词组成的语词。

6.1 语词形式

注：对于汉语，表达概念的语词可以是单字(如：跑、铁、春)、多字词(如：微波、拖拉机、解放战争)或短语(如：中国特色社会主义、可持续发展战略)。

——在“符号、简称和其他约定”中，增加了常用汉语标记符：

汉语标记符
Y 或用
D 或代
S 或属
F 或分
C 或参

——在“6 叙词表语词”中，增加了有别于英语的汉语语词语法形式、标点符号、特殊字符等，集中编列。例如：增加了“6.4.3 汉语语词的符号”“6.6.9.4 专有名称的其他类型”。

——在“7 复杂概念”中，增加了“7.2.2 汉语复合语词”的内容，与英语复合词分别编列。

——在“9 跨语言等同”中，只保留了与英语、德语、法语相关的实例，其他语言(如西班牙语、荷兰语等)则省略。

——在“12 显示与布局”中，增加了汉语叙词表的示例图，如：

图 6 叙词表语词字顺序显示(汉语叙词表)

图 8 叙词表语词按分面组织的等级显示(汉语叙词表)

图 9 叙词表语词按主题领域排列的分类显示

图 11 叙词表语词按主题领域分面排列的分类显示

图 14 双语种叙词表中单条记录显示(汉语-英语)

图 15 双语种叙词表中文版字顺表显示

等等。

——在“附录”中,增加了汉语叙词表显示实例,编号为附录 B。将国际标准的附录 B 改为附录 C。

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这些专利的责任。

本部分由全国信息与文献标准化技术委员会(SAC/TC 4)提出并归口。

本部分主要起草单位:华东理工大学、国家图书馆、中国科技信息研究所、北京大学、解放军南京政治学院、中国国防科技信息中心。

本部分主要起草人:陈树年、汪东波、曾建勋、常春、王军、包冬梅、真溱。

本部分所代替标准的历次版本发布情况为:

——GB/T 13190—1991;

——GB/T 15417—1994。

引 言

由纸本转化而来的现代叙词表也是适用于数字环境的信息组织工具,它们的构建和维护依托软件支持并能够实现与搜索引擎及内部管理系统等其他应用程序的集成(叙词表数据库中的数据或许需要和检索应用程序中找到的记录数量结合提供)。以往叙词表主要用于信息专业人员的标引和检索,当今叙词表要适应未受过培训的用户更直观地发现信息需求及其使用,并且要满足通过机器产生新词汇的要求。

虽然叙词表标准需要向信息管理电子化过渡,但是 GB/T 13190 的本部分仍假定在选择标引词和检索词时通常需要人的智能参与。叙词表设计的基本原则是指引标引人员和检索人员对于同一概念选择同样的语词,从而检索到相关文献,即使将叙词表应用于计算机的情况下也是如此。

数据有效交换是叙词表管理和开发的一个重要组成部分。本部分还包括了交换格式和协议的若干建议。采用这些建议将有利于叙词表管理系统之间以及和利用数据的其他计算机应用程序(如标引和检索系统)之间的互操作。

本部分涵盖叙词表的开发与维护,对于涉及多语种论题及其示例的地方,试图尽可能宽泛地选择语种。

叙词表主要用于后组检索系统,但也可以应用于等级目录、先组索引和分类系统。叙词表应用需要和其他系统,如自动分类系统、自由文本检索系统等相匹配。

GB/T 13190 包括两个部分:本部分是用于信息检索的叙词表;另外的部分是叙词表互操作(待制订),涉及其他类型的结构化词表(如专业分类表、规范文档、本体等),并给出了在信息存储和检索过程各阶段能够保证词表互操作的建议。

信息与文献 叙词表及与其他词表的互操作

第 1 部分：用于信息检索的叙词表

1 范围

GB/T 13190 的本部分为用于信息检索的叙词表的开发与维护提供了一个推荐性的建议。该建议不受应用媒体类型(文本、声音、静止或动态图像、物理实体或多媒体)的影响,适用于检索所有类型信息资源,包括知识库与知识门户、书目数据库、文本、博物馆与多媒体馆藏及其各自的词表。

本部分还提供了一个数据模型以及叙词表数据输入输出的推荐格式。

本部分适用于单语种和多语种叙词表。

本部分不适用于书后索引的编制,但其中很多规定对编制书后索引是有用的。

本部分不适用于在检索或标引应用程序中直接使用的数据库或软件,但也预见这类应用程序在对叙词表管理实用性推荐方面的应用需求。

2 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。

GB/T 4880.1 语种名称代码 第 1 部分:2 字母代码 (GB/T 4880.1—2005, ISO 639-1:2002, MOD)

GB/T 4880.2 语种名称代码 第 2 部分:3 字母代码 (GB/T 4880.2—2000, eqv ISO 639-2:1998)

GB/T 7408 数据元和交换格式 信息交换 日期和时间表示法 (GB/T 7408—2005, ISO 8601:2000, IDT)

GB/T 22466 索引编制规则(总则) (GB/T 22466—2008, ISO 999:1996, NEQ)

ISO 12199 多语种术语的字母数字顺序和词典数据的拉丁字母显示 (Alphabetical ordering of multilingual terminological and lexicographical data represented in the Latin alphabet)

ISO 12620 术语和其他语种及内容资源 数据分类规范和语言资源的数据分类注册系统管理 (Terminology and other language and content resources—Specification of data categories and management of a data category registry for language resources)

ISO 16642 计算机在术语中的应用 术语置标框架 (Computer applications in terminology—Terminological markup framework)

ISO 25577 信息与文献 MARC 的 XML 显示 (Information and documentation—MarcXchange)

ISO/IEC 14651 信息技术 国际字符顺序及对照 字符串对照方法和可更改的通用字符顺序模板 (Information technology—International string ordering and comparison—Method for comparing character strings and description of the common template tailorable ordering)

ANSI/NISO Z39.19:2005 单语种受控词表构建、格式和管理指南 (Guidelines for the Construction, Format, and Management of Monolingua Controlled Vocabularies)

ANSI/NISO Z39.50 信息检索:应用服务定义和协议规范 (Information Retrieval: Application Service Definition & Protocol Specification)