



中华人民共和国国家标准

GB/T 19708—2005/ISO 12199:2000

用拉丁字母系统表示的多语种术语和辞书 资料字母排序规则

Alphabetical ordering of multilingual terminological and lexicographical data
represented in the Latin alphabet

(ISO 12199:2000, IDT)

2005-03-23 发布

2005-10-01 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

目 次

前言	III
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 准备工作	2
5 第一级排序	2
6 第二级排序	4
7 第三级排序	5
8 第四级排序	6
附录 A (规范性附录) 逐词排序规则	7
附录 B (资料性附录) 辞典与术语排序的特殊规则	8
附录 C (资料性附录) 化学名称的排序规则	9
附录 D (资料性附录) 拉丁字母系统的字符表	11
附录 E (资料性附录) 使用拉丁字母系统的语种	20
附录 F (资料性附录) 字母顺序和字符表	27
附录 G (规范性附录) 标准规则的规范性说明	40

前 言

本标准等同采用 ISO 12199:2000《用拉丁字母系统表示的多语种术语和辞书资料字母排序规则》(英文版)。

本标准的附录 A、附录 G 为规范性附录,附录 B、附录 C、附录 D、附录 E 和附录 F 为资料性附录。

本标准由全国术语标准化技术委员会提出。

本标准由全国术语标准化技术委员会归口。

本标准由中国标准化研究院起草。

本标准主要起草人:于欣丽、叶盛、陈玉忠、张志云、程永红、徐俊荣、肖玉敬、宋敏、卢丽丽。

引 言

在撰写术语集和辞书资料的工作(包括打印格式和数据库形式)中,为了便于用户的检索,有必要遵循国际公认的用于术语和辞书资料的拉丁字母排序规则。同时,这种规则也有利于术语和辞书资料的交换,因此采用国际标准规定的拉丁字母排序规则是非常必要的。

用拉丁字母系统表示的多语种术语和辞书 资料字母排序规则

1 范围

本标准规定了用拉丁字母表示的多语种术语和辞书资料(术语、术语成分或单词)按字母排序的规则。本标准已经考虑到了已有的术语和辞书资料记录中用拉丁字母表示的各种语言的字符集,同时还考虑到将其他语种的资料按国际规范的约定转写成拉丁字母时所使用的字符集。本标准给出的字母字符规则只适用于同时使用多语种的场合,并不影响各语种本身的字母排序。

本标准的正文规定了对字符串进行逐字母排序的方法。

本标准的附录 A 给出了逐词排序的规则,逐词排序目前已被广泛使用。

本标准的附录 B 介绍了两条补充的规则,有助于辞书和术语的排序。

本标准的附录 C 介绍了化学名称的排序规则。

本标准的附录 D 列出了拉丁字母系统的各种字符。

本标准的附录 E 列出了使用拉丁字母系统的语种。

本标准的附录 F 给出了一些使用拉丁字母系统的语种偏离本标准规定的字母顺序。

本标准的附录 G 对本标准的正文进行了规范性的说明,规则符合 ISO/IEC 14651。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本标准,然而,鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本标准。

GB/T 4880—1991 语种名称代码(eqv ISO 639:1988)

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

GB 13000.1 信息技术 通用多八位编码字符集(UCS) 第一部分:体系结构与基本多文种平面
(GB 13000.1—1993, idt ISO/IEC 10646.1—1993)

GB/T 15237.1—2000 术语工作 词汇 第1部分 理论与应用(eqv ISO 1087-1:2000)

GB/T 17532—1998 术语工作 计算机应用 词汇(eqv ISO/DIS 1087-2-2:1996)

ISO/IEC 14651 信息技术 国际字符串排序约定 字符串的比较方法和按需裁减的排序方法

3 术语和定义

GB/T 15237.1—2000 和 GB/T 17532—1998 确立的以及下列术语和定义适用于本标准。

3.1

字符 character

供组织、控制或表示数据用的元素集合中的一个元素。

3.2

字母 letter

图形字符,可单独或同别的字母组合使用,主要用于代表会话语言中的一个发声单元。

3.3

数字 digit

用于表示数值或数码的字符。