



中华人民共和国国家标准

GB/T 18155—2000
eqv ISO 12200:1999

术语工作 计算机应用 机器可读术语交换格式(MARTIF) 协商交换

Terminology work—Computer applications
—Machine-readable terminology interchange format(MARTIF)
—Negotiated interchange

2000-07-14 发布

2001-03-01 实施

国家质量技术监督局 发布

目 次

前言	Ⅱ
ISO 前言	Ⅳ
引言	V
1 范围	1
2 引用标准	1
3 定义	1
4 结构化术语信息	2
5 MARTIF 的术语条目	3
5.1 数据类目	3
5.1.1 数据类目的规范	3
5.1.2 MARTIF 标记	3
5.1.3 MARTIF 属性	7
5.1.4 类型(<i>type</i>)属性的值	8
5.2 MARTIF 条目结构	9
5.2.1 MARTIF 文件结构	9
5.2.2 术语条目	9
5.2.3 准等义的处理	11
5.2.4 术语条目<termEntry>的使用规则	14
5.2.5 链接	18
6 字符编码和语种(<i>lang</i>)属性	19
7 交换过程	20
8 MARTIF 文件类型定义(DTD)	20
8.1 术语文件的总体结构	20
8.2 前导说明部分	25
8.2.1 前导声明	25
8.2.2 MARTIF 框架	27
8.2.3 MARTIF 正文	27
8.2.4 MARTIF 字符实体	27
8.3 MARTIF 文件头	28
8.4 MARTIF 正文	29
8.4.1 MARTIF 组成部分	29
8.4.2 MARTIF 正文前部	29
8.4.3 MARTIF 正文主体	29
8.4.4 MARTIF 正文后部	30
8.5 验证	33

附录 A(标准的附录)	规范数据类目的表达	40
附录 B(提示的附录)	文献题录条目置标	99
附录 C(提示的附录)	按照通用标识符(GI)和属性(attribute)列表的数据类目	108
附录 D(提示的附录)	数据建模的不同形式	116
附录 E(提示的附录)	MARTIF 文件实例	117
附录 F(提示的附录)	来自 GB/T 17532 和 GB/T 14814 的术语和定义	131
附录 G(提示的附录)	使用 GB 2312 的 SGML 声明	131

前 言

本标准等效采用 ISO 12200:1999《术语工作——计算机应用——机器可读术语交换格式 (MARTIF)——协商交换》。在遵循 ISO 12200 技术内容的前提下,本标准对有关汉语术语数据的问题做了明确的规定和说明,主要是:根据汉语汉字以及我国术语数据的特点,规定了交换汉语术语数据时注明引用汉字字符集的方法;分别编写了使用字符集 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》和 GB 2312《信息交换用汉字编码字符集 基本集》的 SGML 声明;增加了有关交换汉语术语数据的说明;用汉语示例代替了部分示例,在一些示例中增加了汉语内容,并对部分示例增加了说明等。

本标准是术语数据库的系列国家标准之一,已经发布的配套标准有:

GB/T 13725—1992 建立术语数据库的一般原则与方法

GB/T 13726—1992 术语与辞书条目的记录交换用磁带格式

GB/T 15387.1—1994 术语数据库开发文件编制指南

GB/T 15387.2—1994 术语数据库开发指南

GB/T 15625—1995 术语数据库技术评价指南

GB/T 16785—1997 术语工作 概念与术语的协调

GB/T 16786—1997 术语工作 计算机应用 数据类目

GB/T 17532—1998 术语工作 计算机应用 词汇

本标准的附录 A 是标准的附录,附录 B、附录 C、附录 D、附录 E、附录 F 和附录 G 都是提示的附录。

本标准由全国术语标准化技术委员会提出。

本标准由中国标准研究中心归口。

本标准由中国标准研究中心、国家语言文字工作委员会、中国科技信息情报研究所、北京大学计算机研究所、大百科全书出版社等单位起草。

本标准由全国术语标准化技术委员会计算机辅助术语工作分技术委员会负责解释。

本标准主要起草人:李莉、于欣丽、冯志伟、周智佑、杨建武、田野、肖玉敬、徐俊荣。

ISO 前言

国际标准化组织(ISO)是各国家标准化团体(ISO 成员团体)的一个世界范围的联盟。国际标准的制定工作通常是通过 ISO 的技术委员会进行的,对已成立了技术委员会的某个专业领域感兴趣的成员团体都有权参加该技术委员会。与 ISO 有联系的官方和非官方的国际组织也参与这项工作。ISO 在电工技术标准化方面与国际电工委员会(IEC)紧密合作。

技术委员会所采纳的国际标准草案需分发给各成员团体投票表决,作为国际标准发布时要求至少 75% 的成员团体投票批准。

国际标准 ISO 12200 是由 ISO/TC 37/SC 3“术语工作(原则与协调)技术委员会第 3 分委员会——计算机辅助术语工作分委员会”制定的。

本国际标准中的规范是在与文本编码倡议协会(TEI)和工业标准本地化协会(LISA)的紧密合作下进行的。

ISO 12200 以 ISO 8879 标准通用置标语言(SGML)为基础,解决协商交换问题,为了覆盖术语数据库和专业词典中出现的术语条目的所有类型和形式,本标准尽量设计成更具开放性和灵活性。

在以后制定的 ISO 12200 其他部分中将规定具体目的下的更为详细的交换格式。即将制定的这些部分的目的是在没有干预的情况下能够在系统间交换更多的数据。根据 ISO 12200 这些未来部分制定的格式同本标准兼容,这样按照这些未来部分建立的文件就可以用根据本标准规定的文件类型声明(DTD)来分析,而按照本标准建立的文件未必能用根据未来部分规定的 DTD 来分析。

对于一般的术语工作,下列国际标准是有关的:ISO 704、ISO 860、ISO 1087 和 ISO 10241。

附录 A 是标准的附录,附录 B、附录 C、附录 D、附录 E、附录 F 和附录 G 仅是提示的附录。

引 言

术语数据的收集、管理和存储使用了多种多样的术语数据库系统,从单独用户使用的个人计算机到大公司或政府部门操作的大型术语库系统。不同应用程序、系统和硬件平台间术语数据的交换正变得越来越必要。设计本标准正是为了满足这些需求以利于高效的数据交换。

GB/T 14814《信息处理 文本和办公系统》标准通用置标语言(SGML)》,提供了一种描述文件的方法。它不是描述在页面上提交文件的编码方法,而是描述文件的结构特点和文件的组成部分间的相互联系。众所周知,SGML 提供了唯一的通用描述语言,许多置标系统用这种语言易于从一个程序或系统传输文本(或信息)到另一个程序或系统。随着 SGML 语言应用的增长,它的使用范围不断扩大,这与设计者的初衷是一致的,即对文本置标以利于数据交换和信息检索,以及对文本编码以便在超文本环境下操作。

本标准以 SGML 语言为基础,解决机器可读术语数据的协商交换问题,不是规定个别数据库的结构和功能。本标准设计得更具开放性和灵活性,以便覆盖术语数据库和专业词典中,以及各种应用程序、操作系统和硬件平台间的各种术语条目结构。它主要是为能在计算机上存储、读取、检索和操作的术语数据设计的,不局限于任何特殊的软件和硬件配置。

本标准的主要目的是:当程序员或分析员设计术语数据库间数据交换的转出转入软件时,为他们提供指导。本标准中规定的文件类型定义(DTD)允许用通用的 SGML 分析器做部分验证(例如证实文件是否符合 DTD 所规定的结构)。

本标准解决机器可读术语交换格式(MARTIF)协商交换问题,所谓“协商”就是在双方初次交换之前,先就双方数据结构、数据的语种、涉及的数据类目等进行讨论,并对转换程序做适当的调整。

本标准应与 GB/T 16786《术语工作 计算机应用 数据类目》一起使用。

中华人民共和国国家标准

术语工作 计算机应用 机器可读术语交换格式(MARTIF) 协商交换

GB/T 18155—2000
eqv ISO 12200:1999

Terminology work—Computer applications
—Machine-readable terminology interchange format(MARTIF)
—Negotiated interchange

1 范围

本标准规定了机器可读术语交换格式(MARTIF)协商交换部分的结构化术语信息、置标用标记和属性、字符编码与语种属性、MARTIF 术语条目结构和文件类型定义(DTD),以及 MARTIF 文件和结构和交换过程。

本标准适用于机器可读术语数据的交换工作和研制相应的转入转出程序,以及研制按GB/T 13726 编码的数据的转换程序。

2 引用标准

下列标准所包含的条文,通过在本标准中引用而构成为本标准的条文。本标准出版时,所示版本均为有效。所有标准都会被修订,使用本标准的各方应探讨使用下列标准最新版本的可能性。

- GB/T 1988—1998 信息技术 信息交换用七位编码字符集(eqv ISO/IEC 646:1991)
- GB 2312—1980 信息交换用汉字编码字符集 基本集
- GB/T 2659—2000 世界各国和地区名称代码(eqv ISO 3166-1:1997)
- GB/T 4880—1991 语种名称代码(eqv ISO 639:1988)
- GB/T 7408—1994 数据元和交换格式 信息交换 日期和时间表示法(eqv ISO 8601:1988)
- GB 13000.1—1993 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(idt ISO/IEC 10646-1:1993)
- GB/T 13143—1991 情报与文献工作词汇 传统文献(neq ISO 5127-2:1983)
- GB/T 13726—1992 术语与辞书条目的记录交换用磁带格式(neq ISO 6156:1987)
- GB/T 14814—1993 信息处理 文本和办公系统 标准通用置标语言(SGML)
(eqv ISO 8879:1986)
- GB/T 16786—1997 术语工作 计算机应用 数据类目(eqv ISO/DIS 12620:1996)
- GB/T 17532—1998 术语工作 计算机应用 词汇(eqv ISO/DIS 1087-2-2:1996)
- ISO 5127-3a:1981 信息和文献词汇——第 3a 部分——文献和数据的获得、识别和分析
- ISO 12083:1994 信息和文献——电子文稿的编制和置标

3 定义

本标准采用 GB/T 14814、GB/T 17532 和 GB/T 16786 中的术语和下列定义: