

ICS 35.240.99

L 82.2



# 中华人民共和国国家标准

GB/T×××—2008/ ISO/IEC 24703:2004

---

## 信息技术 学习、教育和培训 参与者标识符

Information technology - Learning, education and training -

Participant identifiers

(ISO/IEC 24703:2004, IDT)

(报批稿)

××××-××-××发布

××××-××-××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.....	II
引 言.....	III
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 一致性.....	2
5 功能性.....	2
6 概念模型.....	2
7 语义.....	2
7.1 参与者标识符的组成.....	2
7.2 数据类型.....	2
7.3 可用的抽象字符.....	3
7.4 编码字符.....	4
8 绑定.....	4
8.1 数据实例.....	4
8.2 C、C++绑定.....	4
8.3 Java 绑定.....	4
8.4 ECMA 脚本（JavaScript）绑定.....	5
9 编码.....	5

## 前 言

本标准等同采用ISO/IEC 24703:2004(E)《信息技术 参与者标识符》。

本标准由教育部提出。

本标准由全国信息技术标准化委员会归口。

本标准起草单位：华南理工大学。

本标准主要起草人：奚建清、陈天、文志康、张杨、郑莉、刘银涛。

## 引言

本规范定义一个标识符的数据类型，用于标识学习、教育和培训过程中的参与者。参与者可以包括系统用户、教师、代理、小组、组织、机构等。



# 信息技术 学习、教育和培训 参与者标识符

## 1 范围

本标准定义学习、教育、培训等领域参与者标识符的数据类型。本标准不包含与使用参与者标识符相关的个人信息安全和隐私保护等内容。本标准也不包含参与者标识符的命名策略、注册和鉴别。

注：使用参与者标识符容易导致对个人隐私信息不恰当的访问和错用。实现者需要保证对参与者标识符的正确使用。

## 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 5271.1—2000 信息技术 词汇 第1部分：基本术语（eqv ISO/IEC 2382-1：1993）

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 18221—2000 信息技术 程序设计语言、环境与系统软件接口 独立于语言的数据类型（idt ISO/IEC 11404：1996）

ISO/IEC TR 10176：2003 信息技术 编程语言标准的制定指南

## 3 术语和定义

GB/T 5271.1—2000 确立的以及下列术语和定义适用于本标准。

### 3.1

**抽象字符** abstract character

GB 13000.1—1993字汇（UCS）中的字符，与编码无关。

### 3.2

**绑定** binding

从一个框架或规范到另一个框架或规范的应用或映射。

### 3.3

**数据类型** datatype

由不同的值组成的集合的一个属性，该属性描述这些不同的值所具有的共同特点以及所能施加的操作。

### 3.4

**编码** encoding

信息的位或字节格式及其表示法。

### 3.5

**实现定义的行为** implementation-defined behaviour

未被规定的行为，这些行为需要在实现时做出选择。

### 3.6

**参与者标识符** participant identifier

与参与者相关联的标识。

注：一个参与者可能拥有多于一个的参与者标识符——一个非单一的标识符，但本标准不定义标识符的单一性策略问题。

### 3.7

#### 参与者信息 participant information

关于信息技术系统中的参与者或与之相关联的信息。

### 3.8

#### 最低峰值 (SPM) smallest permitted maximum

应用程序至少能支持的列表项数或字符串长度等。

例如：字段X的最低峰值长度应为25。

### 3.9

#### 通用多八位编码字符集 (UCS) universal character set

由GB 13000.1—1993定义的字符集

### 3.10

#### 不确定行为 unspecified behaviour

一个标准的执行行为提供了两种或多种可能性，并且没有在任何实例中建议哪种行为会得到采用。

## 4 一致性

一致的参与者标识符应该符合第5章、第6章、第7章，同时符合第8章中至少一种绑定，并且符合第9章中规定的要求。

## 5 功能性

参与者标识符应该表示一个指示符。

注：一个参与者可以与一个以上的参与者标识符相关联。

参与者标识符应该是可打印的文本串，该文本串应该便于人类抄录。

示例：

长度为10个字符的数字字符串比混有拉丁字符的长度为256个字符的字符串更容易抄写。

## 6 概念模型

参与者标识符是由抽象字符构成的和一个参与者相关联的字符串。概念上，参与者标识符是“不透明的”，也就是说字符串本身不包含任何意义，如：“35267901”没有任何特定含义。但是，某些实现者和管理者会选择能够传递某种信息的参与者标识符，如：“张三”。一些参与者标识符甚至包含了更多的信息，例如“二楼的张三”。

一个参与者标识符可以用于若干种情境，例如可以被嵌入文件名、URL或者电子邮件地址。在一个管理参与者信息的信息技术系统中，参与者标识符可以被用于连接不同的库，来提供普通的和（或）统一的标识方法。

## 7 语义

### 7.1 参与者标识符的组成

参与者标识符由多个标识符段和分隔符组成。标识符段是由抽象字符集中的抽象字符组成。分隔符在实现时定义，建议采用“.”。

### 7.2 数据类型

下面的数据类型用于定义参与者标识符，该数据类型来自GB/T 18221—2000：

characterstring (GB 13000.1—1993)

参与者标识符长度的最低峰值 (SPM) 是 50个抽象字符。

注：参数“GB 13000.1—1993”标识了抽象字符库，不是编码。

### 7.3 可用的抽象字符

参与者标识符中的抽象字符应局限于以下集合。其中的抽象字符均以UCS中的基本多语种平面中的16位十六进制名称来命名，这些抽象字符的子集是基于ISO/IEC TR 10176: 2003标准的。

#### 7.3.1 标识符字符集

下列是非数字，非特殊字符的抽象字符集

- 拉丁文 (Latin) : 0041~005A、0061~007A、00AA、00BA、00C0~00D6、00D8~00F6、00F8~01F5、01FA~0217、0250~02A8、1E00~1E9B、1EA0~1EF9、207F
- 希腊文 (Greek) : 0386、0388~038A、038C、038E~03A1、03A3~03CE、03D0~03D6、03DA、03DC、03DE、03E0、03E2~03F3、1F00~1F15、1F18~1F1D、1F20~1F45、1F48~1F4D、1F50~1F57、1F59、1F5B、1F5D、1F5F~1F7D、1F80~1FB4、1FB6~1FBC、1FC2~1FC4、1FC6~1FCC、1FD0~1FD3、1FD6~1FDB、1FE0~1FEC、1FF2~1FF4、1FF6~1FFC
- 西里尔文 (Cyrillic) : 0401~040C、040E~044F、0451~045C、045E~0481、0490~04C4、04C7~04C8、04CB~04CC、04D0~04EB、04EE~04F5、04F8~04F9
- 亚美尼亚文 (Armenian) : 0531~0556、0561~0587
- 希伯来文 (Hebrew) : 05B0~05B9、05BB~05BD、05BF、05C1~05C2、05D0~05EA、05F0~05F2
- 阿拉伯文 (Arabic) : 0621~063A、0640~0652、0670~06B7、06BA~06BE、06C0~06CE、06D0~06DC、06E5~06E8、06EA~06ED
- 天成文 (Devanagari) : 0901~0903、0905~0939、093E~094D、0950~0952、0958~0963
- 孟加拉文 (Bengali) : 0981~0983、0985~098C、098F~0990、0993~09A8、09AA~09B0、09B2、09B6~09B9、09BE~09C4、09C7~09C8、09CB~09CD、09DC~09DD、09DF~09E3、09F0~09F1
- 锡克教文 (Gurmukh) : 0A02、0A05~0A0A、0A0F~0A10、0A13~0A28、0A2A~0A30、0A32~0A33、0A35~0A36、0A38~0A39、0A3E~0A42、0A47~0A48、0A4B~0A4D、0A59~0A5C、0A5E、0A74
- 古吉特拉文 (Gujarati) : 0A81~0A83、0A85~0A8B、0A8D、0A8F~0A91、0A93~0AA8、0AAA~0AB0、0AB2~0AB3、0AB5~0AB9、0ABD~0AC5、0AC7~0AC9、0ACB~0ACD、0AD0、0AE0
- 奥利雅文 (Oriya) : 0B01~0B03、0B05~0B0C、0B0F~0B10、0B13~0B28、0B2A~0B30、0B32~0B33、0B36~0B39、0B3E~0B43、0B47~0B48、0B4B~0B4D、0B5C~0B5D、0B5F~0B61
- 泰米尔文 (Tamil) : 0B82~0B83、0B85~0B8A、0B8E~0B90、0B92~0B95、0B99~0B9A、0B9C、0B9E~0B9F、0BA3~0BA4、0BA8~0BAA、0BAE~0BB5、0BB7~0BB9、0BBE~0BC2、0BC6~0BC8、0BCA~0BCD
- 泰卢固文 (Telugu) : 0C01~0C03、0C05~0C0C、0C0E~0C10、0C12~0C28、0C2A~0C33、0C35~0C39、0C3E~0C44、0C46~0C48、0C4A~0C4D、0C60~0C61
- 卡纳达文 (Kannada) : 0C82~0C83、0C85~0C8C、0C8E~0C90、0C92~0CA8、0CAA~0CB3、0CB5~0CB9、0CBE~0CC4、0CC6~0CC8、0CCA~0CCD、0CDE、0CE0~0CE1
- 德拉维族文 (Malayalam) : 0D02~0D03、0D05~0D0C、0D0E~0D10、0D12~0D28、0D2A~0D39、0D3E~0D43、0D46~0D48、0D4A~0D4D、0D60~0D61
- 泰文 (Thai) : 0E01~0E3A、0E40~0E5B
- 老挝文 (Lao) : 0E81~0E82、0E84、0E87~0E88、0E8A、0E8D、0E94~0E97、0E99~0E9F、0EA1~0EA3、0EA5、0EA7、0EAA~0EAB、0EAD~0EAE、0EB0~0EB9、0EBB~0EBD、0ECO~0EC4、0EC6、0EC8~0ECD、0EDC~0EDD
- 藏文 (Tibetan) : 0F00、0F18~0F19、0F35、0F37、0F39、0F3E~0F47、0F49~0F69、0F71~0F84、0F86~0F8B、0F90~0F95、0F97、0F99~0FAD、0FB1~0FB7、0FB9
- 格鲁吉亚文 (Georgian) : 10A0~10C5、10D0~10F6

- 平假名 (Hiragana) : 3041~3093、309B~309C
- 片假名 (Katakana) : 30A1~30F6、30FB~30FC
- 注音 (Bopomofo) : 3105~312C
- CJK 统一汉字 (CJK Unified Ideographs) : 4E00~9FA5
- 韩文 (Hangul) : AC00~D7A3

注：“~”表示连续，如“0FB1~0FB7”说明从0FB1起的连续七个代码表示了七个藏文字符。以下“~”的含义类同。

### 7.3.2 数字

下列是数字字符：

0030~0039、0660~0669、06F0~06F9、0966~096F、09E6~09EF、0A66~0A6F、0AE6~0AEF、0B66~0B6F、0BE7~0BEF、0C66~0C6F、0CE6~0CEF、0D66~0D6F、0E50~0E59、0ED0~0ED9、0F20~0F33

注：“~”表示连续。

### 7.3.3 特殊字符

下列是特殊字符：

005F、00B5、00B7、02B0~02B8、02BB、02BD~02C1、02D0~02D1、02E0~02E4、037A、0559、093D、0B3D、1FBE、203F~2040、2102、2107、210A~2113、2115、2118~211D、2124、2126、2128、212A~2131、2133~2138、2160~2182、3005~3007、3021~3029

注：“~”表示连续。

## 7.4 编码字符

下列字符定义了编码字符集：

A B C D E F G H I J K L M  
N O P Q R S T U V W X Y Z  
a b c d e f g h i j k l m  
n o p q r s t u v w x y z  
0 1 2 3 4 5 6 7 8 9 \_ -

注1：字符的选择和 IETF RFC 2396 相一致

注2：百分号“%”被用于参与者标识符的编码，但不是“编码字符集”的一部分。

## 8 绑定

### 8.1 数据实例

参与者标识符的实例应该被编码为八位位组的数组。

### 8.2 C、C++绑定

C/C++语言翻译系统下声明兼容的应用程序或环境应该在结构、指针、数组和函数参数中使用“char”基本类型（或相兼容的）来存储参与者标识符。

注：如果7位编码被用于C/C++翻译环境，最低峰值要求隐含说明参与者标识符至少是501个字符长（“%U”+8个16进制数）\*50+1个空终结符。

参与者标识符应该被编码为八位字节的数组。

### 8.3 Java 绑定

声称与Java语言转换系统相一致的应用和环境应该在类、数组和函数参数中使用“String”类型（或相兼容的）来存储参与者标识符。

#### 8.4 ECMA 脚本（JavaScript）绑定

声称与ECMAScript（JavaScript）语言转换系统相一致的应用和环境应该在数组和函数参数中使用“String”类型（或相兼容的）来存储参与者标识符。

## 9 编码

所有的绑定都应该使用本章所规定的编码。

注 1：应用和用户接口可以允许使用对人来说更自然的形式录入和显示这些参与者标识符。如：一个日本用户的登录屏幕上应该只显示 Kanji 字符，同时，应用程序将 Kanji 字符和参与者标识符编码格式进行双向转换。

对每一个抽象字符进行编码时，每个参与者标识符要遵守下面对编码字符的约定中的一条。下面用符号“h”来表示任意一个 16 进制数字（0~9、A~F、大小写不敏感）。

——抽象字符属于编码字符集时，可以直接使用它们自己进行编码。

——某个抽象字符集在 UCS 中 0000 到 00FF 的范围内时，对其中的抽象字符可以采用下面的形式编码：

`%hh`

其中“h”表示单个十六进制字符。

——某个抽象字符集在 UCS 中 0000 到 FFFF 的范围内时，对其中的抽象字符可以采用下面的形式编码：

`%uhhhh`

其中“h”表示单个十六进制字符，“u”是小写字母。

——某个抽象字符集在 UCS 中 00000000 到 FFFFFFFF 的范围内时，对其中的抽象字符可以使用以下形式编码：

`%Uhhhhhhh`

其中“h”表示单个十六进制字符，“U”是大写字母。

注 2：这些编码的约定是与现存的 C、C++、Java 和 IETF 的标准和规范相一致的。

注 3：该编码方案允许一个参与者标识符在 UCS 字符指令集的基础上，以 7 位二进制的形式进行编码。为了提高效率，不同的系统会采用不同的字符集。例如，当 UCS-2 用于数据交换时，Unicode 字符就可以直接使用，不必转化成“%uhhhh”的形式。为了避免丢失符号“%”，在不同的实现中，会有不同的解决方法。

注 4：本标准没有推荐任何一种编码方案，实现者可以选择一种和参与者标识符具体应用相关的编码方案