



# 中华人民共和国国家标准

GB 18030—2022

代替 GB 18030—2005

## 信息技术 中文编码字符集

Information technology—Chinese coded character set

2022-07-19 发布

2023-08-01 实施

国家市场监督管理总局  
国家标准化管理委员会 发布

## 目 次

前言 .....	I
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 字汇 .....	2
5 总体结构 .....	2
6 字符的排列顺序 .....	4
7 码位分配 .....	4
8 部分字符和代码的说明 .....	7
9 实现的级别 .....	7
附录 A (规范性) 双字节字符表 .....	9
附录 B (规范性) 表意文字描述符 .....	91
附录 C (规范性) 四字节字符表 .....	92
附录 D (资料性) 部分字符和代码的说明 .....	546
附录 E (资料性) 《通用规范汉字表》汉字的代码位置 .....	549
参考文献 .....	743

## 前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第 1 部分：标准化文件的结构和起草规则》的规定起草。

本文件代替 GB 18030—2005《信息技术 中文编码字符集》，与 GB 18030—2005 相比，除结构调整和编辑性改动外，主要技术变化如下：

- a) 增加了本文件适用对象(见第 1 章)；
- b) 在双字节编码区更改了 10 个竖排标点和 8 个汉字构件所对应的 GB/T 13000 代码位置，并删除了 6 个重复编码的汉字构件和 9 个重复编码的汉字(见附录 D, 2005 年版的附录 A)；
- c) 在四字节编码区更改了 18 个 GB/T 13000 代码位置(见附录 D, 2005 年版的附录 D)；
- d) 在四字节编码 0x82358F33~0x82359636 部分增加了 CJK 统一汉字新增的 66 个汉字(见附录 C)；
- e) 在四字节编码 0x9835F738~0x98399E36 部分增加了 CJK 统一汉字扩充 C 的 4149 个汉字(见附录 C)；
- f) 在四字节编码 0x98399F38~0x9839B539 部分增加了 CJK 统一汉字扩充 D 的 222 个汉字(见附录 C)；
- g) 在四字节编码 0x9839B632~0x9933FE33 部分增加了 CJK 统一汉字扩充 E 的 5762 个汉字(见附录 C)；
- h) 在四字节编码 0x99348138~0x9939F730 部分增加了 CJK 统一汉字扩充 F 的 7473 个汉字(见附录 C)；
- i) 在四字节编码 0x81398B32~0x8139A035 部分增加了康熙部首 214 个(见附录 C)；
- j) 在四字节编码 0x8134F932~0x81358437 部分增加了 83 个西双版纳新傣文字符(见附录 C)；
- k) 在四字节编码 0x81358B32~0x81359933 部分增加了 127 个西双版纳老傣文字符(见附录 C)；
- l) 在四字节编码 0x82369535~0x82369A32 部分增加了 48 个傣僰文字符(见附录 C)；
- m) 在四字节编码 0x9034C538~0x9034C730 部分增加了 13 个蒙古文 BIRGA 符号(见附录 C)；
- n) 在四字节编码 0x9232C636~0x9232D635 部分增加了 149 个滇东北苗文字符(见附录 C)；
- o) 删除了附录“追加的汉字及部首/部件”(见 2005 年版的附录 C)；
- p) 增加了“实现的级别”一章(见第 9 章)；
- q) 增加了资料性附录“《通用规范汉字表》汉字的代码位置”(见附录 E)。

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专利的责任。

本文件由中华人民共和国工业和信息化部提出并归口。

本文件及其所代替文件的历次版本发布情况为：

- 2000 年首次发布为 GB 18030—2000, 2005 年第一次修订；
- 本次为第二次修订。

# 信息技术 中文编码字符集

## 1 范围

本文件规定了信息技术用的中文图形字符及其二进制编码的十六进制表示。

本文件适用于中文和其他文字图形字符信息的处理、交换、存储、传输、显现、输入和输出。

本文件适用于为具备中文和其他文字图形字符信息化处理及交换功能的技术类产品,包括但不限于以输入法、光学字符识别(OCR)、编辑校对、机器翻译、语音合成、文字转写、智能写作等为代表的软件产品,以及以计算机、通信终端设备、电子书阅读器、学习机等为代表的硬件产品。

## 2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;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 2312—1980 信息交换用汉字编码字符集 基本集

GB/T 11383—1989 信息处理 信息交换用八位代码结构和编码规则

GB/T 13000 信息技术 通用多八位编码字符集(UCS)

## 3 术语和定义

下列术语和定义适用于本文件。

### 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

**编码字符 coded character**

字符(3.1)及其编码表示。

### 3.3

**用户自定义区 private use area**

使用符合本文件的产品的使用者可以自行规定的区域。

### 3.4

**字汇 repertoire**

用编码字符(3.2)集表示的一个指定的字符(3.1)集合。

### 3.5

**保留区 reserved zone**

留作未来本文件规定的区域。