

国家标准《信息技术 中文编码字符集》（征求意见稿）

编制说明

2018-07

一、工作简况

[内容包括下达计划任务主管部门的完整名称、项目计划发布文件号、本项目的计划代号和主要承办单位完整名称、副主办单位或协作单位完整名称、主要工作过程、主要起草人及其所做的工作等]

根据国家标准化管理委员会下达的2017年第XX批国家标准计划任务，《信息技术 中文编码字符集》（20171716-Q-469）由中国电子技术标准化研究院起草。

本标准立项前，中国电子技术标准化研究院自2015年就已组织业内部分单位的有关技术人员对本标准的修订方案进行了预研，确定了主要原则。此后，至2017年，多次以邮件、电话、会议等方式讨论了技术细节。

2017年12月，在接到根据国家标准化管理委员会下达的任务后，正式成立了由北京北大方正电子有限公司、北京亚细亚智业科技有限公司、微软（中国）公司等单位的有关技术人员标准修订起草组，全面启动修订工作。

2015年：草案第一稿

2017年10月：草案第二稿

2018年1月：草案第三稿

2018年4月：草稿第四稿

2018年5月：草稿第五稿

2018年6月，内审稿

2018年7月，标准起草组邀请部分业内专家对本标准草稿进行了审查和会议研讨。会后，标准起草组根据专家的意见对有关内容进行了研究，对标准文本作了相应修改，形成了本标准的征求意见稿。

二、标准编制原则和确定主要内容的论据及解决的主要问题

[如技术指标、参数、公式、性能要求、试验方法、检验规则等的论据，包括试验、统计数据，解决的主要问题。修订标准时应列出与原标准的主要差异和水平对比]

标准工作组确定以下编制原则：

- 1、本标准的结构与上一版（GB 18030-2005）保持一致。
- 2、本次修订仅收录中国各民族现行文字及其他常用文字、标点、符号如拉丁字母、阿拉伯数字等。

2.1、保留上一版收录的全部文字、标点、符号，但其中的维吾尔文、哈萨克文、柯尔克孜文收录情况须根据GB 21669-2008《信息技术 维吾尔文、哈萨克文、柯尔克孜文编码字符集》的规定进行修订。（该标准已于2017年转为推荐性质。）

2.2、收录2013年国务院批准发布的《通用规范汉字表》规定的全部8105个汉字。

2.3、根据GB 13000（或ISO/IEC 10646）最新版情况补充下列文字（其中部分已经收录于本标准的上一版）：

- ..CJK统一汉字及其扩充A、B、C、D、E、F
- ..藏文
- ..规范彝文
- ..朝鲜文
- ..维吾尔文、哈萨克文、柯尔克孜文
- ..蒙古文（包括传统蒙古文、满文、锡伯文、托忒文、阿礼嘎礼字）
- ..滇东北苗文
- ..傣傣文
- ..傣文（共三种：西双版纳新傣文、西双版纳老傣文、德宏傣文）

- 3、本标准规定的编码与GB 13000（idt. ISO/IEC 10646）规定的编码一一对应。

本标准上一版规定的GB 13000临时码位，应根据相应字符在GB 13000（或ISO/IEC 10646）最新版中的编码情况进行调整。

三、主要试验[或验证]情况分析

本标准规定的全部内容已经操作系统及相关文字处理产品验证可行。

四、知识产权情况说明

[相关知识产权情况的说明。如果在标准编制过程中识别出标准的某些技术内容涉及专利，则应列出相关专利的目录及其使用理由]

无。

五、产业化情况、推广应用论证和预期达到的经济效果

[电子行业标准必须填写。对于国家标准如不要求此内容可删除章号和标题]

预期本标准发布实施后，可取得下面效果：

- 1、有利于促进包括现代汉字和少数民族文字在内的中文在信息处理系统中的实现。
- 2、有利于统一文字编码，实现多文种统一处理。
- 3、有利于促进2013年国务院批准发布的《通用规范汉字表》在信息技术领域的落实。

六、采用国际标准和国外先进标准情况

[采用国际标准和国外先进标准的程度，以及与国际、国外同类标准水平的对比情况，国内外关键指标对比分析或与测试的国外样品、样机的相关数据对比情况]

本标准为我国独创，按照本标准开发的软件可以顺利兼容我国现有的按照GB 2312开发的软件，因此适用于我国。按照本标准开发的软件可通过标准提供的代码映射表实现与国际标准ISO/IEC 10646的代码转换，解决国内外信息交流问题。

七、与现行相关法律、法规、规章及相关标准的协调性

[与相关的现行法律、法规和规章的符合性，特别是与强制性国家标准的协调性，以及与同类标准和标准体系中其他标准的协调性说明]

本标准与我国的现行法律、法规和强制性标准无冲突。

八、重大分歧意见的处理经过和依据

本标准的总体结构不作调整，实现技术已经成熟，因而在研制过程中无重要技术分歧。

九、标准性质的建议

[建议是强制性标准还是推荐性标准，对于强制性标准必须列出强制的理由，强制范围]

本标准上一版是当前使用中最重要的中文编码字符集，是我国全社会汉字信息处理能力的基本保障，经2017年国家标准化委员会清理整顿强制性国家标准后仍然保持了强制性质。因此，建议本标准在本次修订后仍然作为条文强制性国家标准发布。对于一般具有文字处理功能的信息技术产品，强制范围与上一版一致，即单字节编码部分、双字节编码部分和从GB 13000（或ISO/IEC 10646）引入的“CJK统一汉字”及其扩充A。

十、贯彻标准的要求和措施建议

[内容包括组织措施、技术措施、过渡办法、实施日期等]

建议本标准发布12个月后实施。

对本标准发布前已经上市并符合GB 18030-2005要求的产品不增加新的强制要求。

十一、替代或废止现行相关标准的建议

[说明废止和/或替代相关标准的情况。报批国家标准时，如有相应转换、代替、废止的电子行业标准或指导性技术文件，亦应说明]

本标准代替GB 18030-2005《信息技术 中文编码字符集》。

十二、其它应予说明的事项

本标准删除了GB 18030-2005的附录C（规范性附录）《追加的汉字及部首/部件》。（在GB 18030-2005中，附录C称“收录了没有在GB 13000.1-1993中收录的汉字及部首/构件”，共计80个。）

本版第4章《字汇》所云来自GB/T 13000的汉字（及部首、构件等）和中国少数民族文字所用字符，多数来自国际标准ISO/IEC 10646: 2017（未来将转化为GB/T 13000），超出GB/T 13000-2010（即GB 13000.1-1993的修订版）的字符数以万计，这些字符的字形及其代码位置已经在本标准的附录A（规范性附录）《双字节字符表》和附录C（规范性附录）《四字节字符表》（GB 18030-2005的附录D）中详细规定。如果效仿GB 18030-2005保留附录《追加的汉字及部首/

部件》而将这些尚未收入GB/T 13000的字符逐一列出，则既无必要又显冗长，因而予以删除。