

国家标准《信息技术 托忒文名义字符、变形显现字符和控制字符使用规则》（征求意见稿）编制说明

一、工作简况

1. 任务来源

根据国家标准化管理委员会 2014 年国家标准制修订计划下达的任务，国家标准《信息技术 托忒文名义字符、变形显现字符和控制字符使用规则》（项目计划代号为 20140222-Q-469）由国家标准化管理委员会提出，全国信息技术标准化技术委员会归口。

该标准由中国电子技术标准化研究院主办，参与起草单位包括内蒙古大学、新疆维吾尔自治区民语委蒙语办、北京北大方正电子有限公司、内蒙古蒙科立软件有限责任公司、新疆大学信息科学与工程学院等。

2. 主要工作过程

- 1) 2014 年 10 月，国家标准化管理委员会正式下达国家标准《信息技术 托忒文名义字符、变形显现字符和控制字符使用规则》（项目计划代号 20140222-Q-469）研制任务；
- 2) 2014 年 11 月 4 日至 5 日，蒙古文信息技术国家标准工作组在内蒙古呼和浩特市组织召开了工作会议，决定国家标准《信息技术 托忒文名义字符、变形显现字符和控制字符使用规则》的研制工作由确精扎布教授负责，并成立了标准编制工作组。
- 3) 2015 年上半年，标准编制组根据多次组织开会研讨，对标准的主要内容、规定范围以及相关技术内容进行了前期的研究和论证。并依据 GB 13000-2010 “信息技术 通用多八位编码字符集 (UCS) (ISO/IEC 10646:2003, IDT)”中的蒙古文编码字符集中规定的托忒文字符，初步起草了标准草案；
- 4) 2015 年 7 月初开始，工作组以最新托忒文字符集、变形显现字符集和“托忒文名义字符到变形显现字符的转换规则”为基础重新整理了标准草案。并于 10 月提交全国信息技术标准化技术委员会秘书处请求内审。

二、 标准编制原则和确定主要内容的论据

本标准规定了托忒文的名义字符、变形显现字符的最小集合，并对托忒文的每个名义字符的变形显现形式的图形、托忒文控制字符的使用规则作出了规定，并规定了托忒文名义字符到变形显现字符的转换规则。本标准适用于托忒文信息处理系统。

1. 编制原则

为了保证本标准的研制工作开展顺利，工作组遵循了以下原则：

- 1) 依据现行国家标准，开展标准研制工作，确保本标准与现行政策、法规、标准的依从关系；
- 2) 充分利用现有学术成果，促进先进学术成果的标准化转化，确保标准的技术水平；
- 3) 注重标准的研制及其应用试验的结合，形成互动，更好地满足标准的推广应用需求。

2. 确定主要内容的依据

分析现行各种托忒文编码和录入法的特点。目前，人们普遍使用的托忒文编码和录入法是方正编码、蒙科立编码、赛音编码、华光编码、蒙古国 Mongolian Editor、微软蒙古文白体。方正编码：从早期的方正出版系统开始到现在的方正 10.0 中都实现了托忒文的处理。目前现行的托忒文编码中只有方正的编码几乎覆盖了托忒文的所有字符。方正编码中反应托忒文的基本字符，阿礼嘎礼字符，数字和标点符号。只缺托忒文特有的标点符号的变体。字体方面，方正实现了托忒文混合字体和独有字体。

蒙科立编码：利用 menk2008 可以录入托忒文。蒙科立编码中只包含托忒文的基本字符部分。该编码中缺乏托忒文的阿礼嘎礼字符，托忒文特有标点符号和数字。蒙科立编码不符合蒙古文编码国际标准，编码中只反映了新疆托忒文的部分字符。

赛因编码：此编码中只反映了托忒文的基本字符部分，缺乏托忒文阿礼嘎礼，托忒文特有标点符号。

华光编码：此编码中包含了托忒文的基本字符，阿礼嘎礼和标点符号。

蒙古国 Mongolian Editor: Mongolian Editor 是西里尔文，传统蒙古文，托忒文混合编辑的记事本。Mongolian Editor 编码中只处理基本托忒文和数字，缺乏托忒文阿礼嘎礼，托忒文特有标点符号。

微软蒙古文白体：这是微软公司基于 UCS 的规定研发的蒙古文字体。在 Windows Vista 以上的系统中都自带的字体，并且在 USP10.dll 中正确解释蒙古文。微软蒙古文白体中包含传统蒙古文，托忒文，满文，锡伯文，阿礼嘎礼，数字，标点符号和控制符。该字体中缺乏托忒文特有的标点符号，变形显现字符和阿礼嘎礼字符。

本托忒文编码国家标准规则是根据《ISO/IEC 10646: 1-2000 (UCS)》的蒙古文字块(1800-18AF)以及有关蒙古文的规定和 1999 年中蒙等国协商通过的《蒙古文编码系统实现用户协定》以及《信息技术 蒙古文表型显现字符集和控制字符使用规则(GB/T 26226-2010)》、《信息技术 信息处理交换用蒙古文字符键盘布局(GB/T 28038-2011)》、《信息技术 托忒文变形显现字符集和控制字符使用规则》(第五稿)，方正蒙文书版 10.0 托忒文录入规则等相关标准编写。

3. 解决的主要问题

目前，研制的托忒文编码国家标准草稿满足以下情况：

- 1) 本托忒文编码遵照国际标准和国家标准的相关规则；
- 2) 全面覆盖托忒文的基本字符，阿礼嘎礼，数字，通用符号和托忒文特有符号；
- 3) 既要满足古托忒文拼写规则，也要满足新疆托忒文的拼写规则。例如：
ᠠᠶᠢᠯᠠᠭᠠᠨ ayil, ᠠᠶᠢᠯᠠᠭᠠᠨ ail, (ᠲᠤᠤᠯᠢ) tuuli, (ᠲᠤᠤᠯᠢ) tuwli, ᠰᠠᠶᠢᠬᠠᠨ sayixan, ᠰᠠᠶᠢᠬᠠᠨ saixan, (ᠲᠤᠤᠵᠢ) tuuji, (ᠲᠤᠤᠵᠢ) tuwji 等；
- 4) 托忒文编码字符由名义字符，单个变形显现字符，强制性合体字和非强制性合体来构成。

三、与现行相关法律、法规、规章及相关标准的协调性

本标准中规定的托忒文字符集是依据 GB 13000-2010 “信息技术 通用多八位编码字符集(UCS) (ISO/IEC 10646:2003, IDT)” 中的蒙古文编码字符集中规定

的托忒文字符。

四、 标准性质的建议

本标准是我国托忒文信息技术标准体系中的重要组成部分，是其中最为重要基础标准，是保证整个标准体系稳定不可或缺的基石。

语言文字是文化的载体，又是文化的结晶，是人们产生民族认同和区别于其他民族的最重要的标志，是重要的民族特征和维系民族内部关系的纽带。虽然满文、锡伯文和托忒文的使用人口不多，在我国只有数十万，但这三种文字特别是满文的影响范围广阔、历史悠久。使用强制标准的方式，可以有效地保证它们得到正确的信息处理。一方面，是满足保存中华民族历史、文化的需要；另一方面，更为重要的是具体落实民族政策，保证少数民族在信息时代能够真正行使法律赋予的使用本民族语言文字的权利。这也是满足当前我们国家进行软实力建设和保持社会稳定、维护民族团结和国家统一的需要。

因此，建议将本标准作为强制性国家标准发布。

五、 贯彻标准的要求和措施建议

为了贯彻实施本国家标准，建议开展本国家标准应用技术的宣传贯彻工作。应在实施前保证文本的充足供应，让每个使用者都能及时得到文本。发布后、实施前应将信息在媒体上进行宣传。对于使用过程中容易出现疑问，可撰文事先予以解释。要分别对标准的不同使用对象，包括出版社、报社、高等院校及其它德宏傣文信息处理应用单位，有侧重点地进行培训、宣传。

2015 年 11 月