



# 中华人民共和国国家标准

GB/T 44596—2024

## 中文域名编码技术要求

Technical specification for coding of Chinese domain names

2024-09-29 发布

2025-01-01 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 .....	III
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语、定义和缩略语 .....	1
3.1 术语和定义 .....	1
3.2 缩略语 .....	3
4 字符串预处理和中文域名字段预处理 .....	3
4.1 字符串预处理框架 .....	3
4.2 字符预备处理概要 .....	4
4.3 映射 .....	5
4.4 归一化 .....	5
4.5 禁止输出 .....	6
4.6 双向字符 .....	6
4.7 字符串预处理框架中的未分配码位 .....	6
4.8 中文域名字段预处理 .....	7
5 Punycode 编码 .....	7
5.1 Punycode 特点 .....	7
5.2 Bootstring 架构 .....	8
5.3 Bootstring 参数 .....	9
5.4 Punycode 参数值 .....	9
5.5 Bootstring 算法 .....	10
附录 A (规范性) UNICOD 保留字符表 .....	13
附录 B (规范性) 映射表 .....	14
附录 C (规范性) 禁止输出表 .....	15
附录 D (规范性) 双向字符表 .....	17
附录 E (资料性) Punycode 编码解码例子 .....	18
附录 F (资料性) Punycode 算法编码解码 C 语言实现例子 .....	20

## 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》给出的规定起草。

本文件为中文域名技术系列标准之一，结构如下：

- 中文域名总体技术要求；
- 中文域名编码技术要求；
- 中文域名解析技术要求；
- 中文域名注册技术要求；
- 中文域名字表技术要求。

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。

本文件由全国通信标准化技术委员会(SAC/TC 485)归口。

本文件起草单位：中国互联网络信息中心、暨南大学、中国科学院计算机网络信息中心、广东盈世计算机科技有限公司、中国信息通信研究院、清华大学、中国电信股份有限公司、中国联合网络通信集团有限公司、中国移动通信集团有限公司、中国通信标准化协会。

本文件主要起草人：姚健康、耿光刚、翁健、李洪涛、李彦彪、段海新、张曼、周琳琳、董科军、杨学、延志伟、孔令飞、沙晓爽、傅瑜、刘越、龙春、赵静、郑晖、王超、孙乐、吴秀诚、张文伟。

# 中文域名编码技术要求

## 1 范围

本文件规定了在互联网体系上使用中文域名的编码规范,以及使用中文域名字段的预处理要求,同时规定了利用国际区域名码进行中文域名编码的技术要求。

本文件适用于各级域名注册管理机构、域名注册服务提供商以及软件厂商开发支持中文域名的应用或者服务等。

## 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。

IETF RFC 3454 国际化字符串预处理(Preparation of internationalized strings)

IETF RFC 3491 国际化域名预处理(Nameprep: a stringprep profile for internationalized domain names)

UNICODE UAX15 通用字符编码标准附录 15(The unicode consortium unicode standard Annex #15)

UNICODE UTR21 通用字符编码技术规范 21(The unicode consortium unicode technical Report #21)

## 3 术语、定义和缩略语

### 3.1 术语和定义

下列术语和定义适用于本文件。

#### 3.1.1

**通用字符编码** **unicode consortium unicode encoding**

根据其位置或码位来识别字符,给每个字符提供的一个唯一的数字。

注:比如,U+4E96指的是在UNICODE字符集中位于4E96处的字符。本文件的UNICODE字符编码采用了ISO/IEC 10646-1,UNICODE字符编码的集合称为UNICODE字符集。

#### 3.1.2

**分隔符** **delimiter**

在punycode里用来分隔的符号“-”。

#### 3.1.3

**域名字段** **domain name label**

域名中由分隔符隔开的几个部分。

注:对于一个完整的域名“www.cnnic.cn”,其中“www”“cnnic”“cn”分别是三个域名字段。