



中华人民共和国国家标准

GB/T 35304—2017

统一内容标签格式规范

Uniform content label format specification

2017-12-29 发布

2018-04-01 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	III
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	2
5 统一内容标签格式描述约定	3
6 统一内容标签格式示意图	4
7 统一内容标签格式定义	5
7.1 UCL 代码部分(UCL Code)	5
7.2 UCL 属性部分(UCL Properties).....	21
7.3 补充说明	39
参考文献	40

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准由全国中文新闻信息标准化技术委员会(SAC/TC 352)提出并归口。

本标准起草单位:东南大学、新华通讯社、国家新闻出版广电总局、清华大学、大有数字资源有限公司、北京大学。

本标准主要起草人:杨鹏、李幼平、吕锐、武国卫、周毅、刁毅刚、尹浩、吕勇强、张长江、郑斌、张利、付蓉、孙宝传、张行功。

引 言

大数据趋势与互联网飞速发展下的媒体变革,已经促使网络空间成为纷繁芜杂的内容大数据的集散地。统一资源定位符 URL(Uniform Resource Locator)是互联网中基础性、核心性和通用性的内容标识标准,内容资源普遍采用 URL 进行标识。但是,以面向地址理念设计的 URL,原理上无法描述内容资源的丰富语义信息,也难以支持基于语义的内容组织和管理,由此带来内容资源难找难管、混乱失序和不可信等问题。这些问题在内容大数据不断涌现的今天显得尤为突出。因此,应对内容大数据带来的挑战,需要在网络空间中引入新的面向内容的标准基石。

本标准针对 URL 的缺陷而提出,旨在全面支持网络空间中内容大数据的共享和治理。统一内容标签 UCL(Uniform Content Label)体现了内容驱动的全新设计理念,它贯穿内容的生产、消费和管理等关键环节,形成三位一体的网络空间内容大数据创新标识体系。统一内容标签可以从多个纬度描述内容资源的特征,支持内容资源的认证、注册及追溯,结构灵活并允许按需裁剪和自由扩展。本标准主要对统一内容标签中各个域的名称、含义、用法及定义规则等格式信息进行描述和规约,为统一内容标签技术的推广、普及和应用,提供关键的标准支持。

统一内容标签格式规范

1 范围

本标准规定了统一内容标签(UCL)的格式,包括统一内容标签中各个域的名称、含义、用法及定义规则等。

本标准适用于对网络空间中面向群体受众的内容资源特征信息进行描述,可支持内容大数据的共享和治理。本标准的使用对象包括通讯社、广播、电视、报刊、新闻网站等内容提供商及软件开发服务商。

注 1: 内容资源的特征信息包括代码信息、语义信息和管理信息等。

注 2: 内容大数据的共享和治理,包括内容资源的聚敛、传播、管理、知识萃取、可信认证和个性化主动服务等。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是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 7408—2005 数据元和交换格式 信息交换 日期和时间表示法

GB/T 13000—2010 信息技术 通用多八位编码字符集(UCS)

GB 18030—2005 信息技术 中文编码字符集

GB/T 20093—2013 中文新闻信息分类与代码

3 术语和定义

下列术语和定义适用于本文件。

3.1

内容元数据 content metadata

反映内容资源特征的描述性信息。

3.2

统一内容标签 uniform content label; UCL

包含代码信息、语义信息和管理信息等的内容元数据。

3.3

统一资源标识符 uniform resource identifier; URI

包含名字或地址的短数据串,指向 web 上的某个对象。

3.4

统一资源定位符 uniform resource locator; URL

包含地址的短数据串,指向 web 上的某个对象。URL 是 URI 的子集。

3.5

包 package

封装整个统一内容标签的数据包。一个 UCL 包由具有一定次序的多个域组成。