



团 体 标 准

T/CAMIR 003—2022

媒体大数据分类分级指南

Guidelines for media big data classification and grading

2022-11-16 发布

2022-11-16 实施

中国市场信息调查业协会 发 布
中 国 标 准 出 版 社 出 版

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 媒体大数据分类分级原则 2

 4.1 多维分类 2

 4.2 多层分级 2

 4.3 就高从严 2

 4.4 自主分类定级 2

5 媒体大数据分类 2

 5.1 分类方法 2

 5.2 分类过程 2

 5.3 分类维度 2

6 媒体大数据分级 4

 6.1 分级方法 4

 6.2 分级要素 4

 6.3 分级规则 5

 6.4 级别变更 6

附录 A（资料性） 媒体大数据分类列表 8

参考文献 11

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。

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。

本文件由中国市场信息调查业协会提出并归口。

本文件起草单位：中国传媒大学、清华大学、广东南方新媒体股份有限公司、杭州凡闻科技有限公司、中科网联数据科技有限公司、北京闻信智远信息科技有限公司、广州数说故事信息科技有限公司、闻客信息(深圳)有限公司、北京易观数智科技股份有限公司、网智天元科技集团股份有限公司、华媒百象传媒科技(北京)有限公司、北京格兰瑞鑫信息科技有限公司、北京蓝象标准咨询服务有限公司。

本文件主要起草人：沈浩、钟茂华、王兵、张燕、李传珍、朱立谷、刘畅、陈法涌、张余、海玲、李旭日、张业军、杨帆、张传文、徐彦超、韩凌、段小莉、马建红、张德保、乔华阳。

媒体大数据分类分级指南

1 范围

本文件提供了媒体大数据分类分级原则、媒体大数据分类和媒体大数据分级的建议。
本文件适用于指导媒体大数据活动中的分类分级。

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 7027—2002 信息分类和编码的基本原则与方法

GB/T 38667—2020 信息技术 大数据 数据分类指南

3 术语和定义

下列术语和定义适用于本文件。

3.1

大数据 big data

具有体量巨大、来源多样、生成极快、且多变等特征,并且难以用传统数据体系结构有效处理的包含大量数据集的数据。

注:国际上,大数据的4个特征普遍不加修饰地直接用 volume、variety、velocity 和 variability 予以表述,并分别赋予了它们在大数据语境下的定义:

- a) 体量 volume:构成大数据的数据集的规模;
- b) 多样性 variety:数据可能来自多个数据仓库、数据领域或多种数据类型;
- c) 速度 velocity:单位时间的数据流量;
- d) 多变性 variability:大数据其他特征,即体量、速度和多样性等特征都处于多变状态。

[来源:GB/T 35295—2017,2.1.1]

3.2

媒体大数据 media big data

以互联网、广播电视网和移动网络等为载体,具有大众传播属性的媒体机构在提供信息服务过程中产生的数据。

3.3

媒体大数据活动 media big data activity

组织针对媒体大数据开展的一组特定任务的集合。

注:主要包括采集、存储、处理、分发和删除等活动。

3.4

大数据分类 big data classification

根据大数据的属性或特征,将其按一定的原则和方法进行区分和归类,并建立起一定的分类体系和排列顺序的过程。

[来源:GB/T 38667—2020,3.3]