



中华人民共和国国家标准

GB/T 20092.2—2022

部分代替 GB/T 20092—2013

中文新闻信息内容 第 2 部分：新闻元数据

Chinese news information content—

Part 2: News metadata

2022-10-12 发布

2022-10-12 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
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 基础公共元数据结构	1
4.1 地点	1
4.2 电子地址	4
4.3 人物	5
4.4 机构	8
4.5 人员组	10
5 文档层元数据	11
5.1 文档层结构	11
5.2 信封	11
6 稿件层元数据	19
6.1 稿件层结构	19
6.2 管理型元数据组	20
6.3 描述型元数据组	28
6.4 权利型元数据组	42
7 内容项层元数据	49
7.1 内容项层结构	49
7.2 内容项物理元数据组	49
7.3 文本内容项元数据组	53
7.4 图片内容项元数据组	55
7.5 图形内容项元数据组	57
7.6 音频内容项元数据组	59
7.7 视音频内容项元数据组	62
7.8 关键帧内容项元数据组	69
7.9 多媒体内容项元数据组	70
7.10 应用文档内容项元数据组	71
附录 A (资料性) 受控词表	73

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。

本文件是 GB/T 20092《中文新闻信息内容》的第 2 部分。GB/T 20092 已经发布了以下部分：

- 第 1 部分：概念模型；
- 第 2 部分：新闻元数据；
- 第 3 部分：数据交换的 XML 格式。

本文件代替 GB/T 20092—2013《中文新闻信息置标语言》中的第 7 章“CNML 公共结构和基本数据类型”、第 11 章“CNML 文档层结构”、第 12 章“CNML 稿件层”、第 13 章“CNML 内容项层”和附录 D，与 GB/T 20092—2013 相比，除结构调整和编辑性改动外，主要技术变化如下：

- a) 按照现有的元数据分组，单独定义元数据组，更改了元数据表的定义格式，保留中文名称、英文名称、说明、类型、属性等内容；
- b) 将“CNML 公共结构和基本数据类型”更改为“基础公共元数据结构”，并将 2013 年版的有关内容更改后纳入（见第 4 章，2013 年版的第 7 章）；
- c) 将“CNML 文档层结构”更改为“文档层元数据”，并将 2013 年版的有关内容更改后纳入（见第 5 章，2013 年版的第 11 章）；
- d) 将“CNML 稿件层”更改为“稿件层元数据”，并将 2013 年版的有关内容更改后纳入（见第 6 章，2013 年版的第 12 章）；
- e) 将“CNML 内容项层”更改为“内容项层元数据”，并将 2013 年版的有关内容更改后纳入（见第 7 章，2013 年版的第 13 章）；
- f) 增加了稿件层元数据的部分元数据定义（见 6.2.11、6.2.12、6.2.13）；
- g) 增加了内容项层元数据的部分元数据定义（见 7.2.10、7.2.11、7.5.5、7.7.8）；
- h) 删除了内容项层元数据的部分元数据定义（见 2013 年版 13.3.4.4.2、13.3.4.4.5、13.3.4.4.14、13.3.4.4.15、13.3.4.5.2、13.3.4.5.5）。

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。

本文件起草单位：新华通讯社、清华大学、国家应用软件产品质量检验检测中心、北京中联合超高清协同技术中心有限公司、北京北大方正电子有限公司、中国科学院计算机网络信息中心、北京智途云天科技有限公司、视觉（中国）文化发展股份有限公司、北京北科光大信息技术股份有限公司、中国人民大学、中国传媒大学、中国大百科全书出版社有限公司、杭州前方信息技术有限公司、东南大学、新华网股份有限公司、中国互联网新闻中心、南方报业传媒集团、重庆日报报业集团、青苹果数据中心有限公司、北京中科大洋科技发展股份有限公司。

本文件主要起草人：武国卫、黄菁、张鹏、王威、王付生、朱建波、朱艳华、丁峰、王骏清、王堃越、许向东、陈锐、张弛、曹学军、刘万福、杨鹏、杨溟、王建平、何崑、张晓伟、张晴、付蓉、路海燕、张贤良。

本文件所代替文件的历次版本发布情况为：

- 2006 年首次发布为 GB/T 20092—2006，2013 年第一次修订；
- 本次为第二次修订。

引 言

GB/T 20092—2013 自发布实施以来,已在很多新闻媒体单位的应用系统中使用。近年来,在标准实施过程中不断涌现一些新的应用场景,产生一些新的应用需求。新闻媒体的部分应用系统也逐渐不再使用 XML 格式,使用 JSON 格式传输稿件的应用越来越多,因为 GB/T 20092—2013 定义的技术内容主要是基于 XML 格式定义的,已经逐渐不适应新闻媒体的应用需求,所以本次修订的主要目的是避免将标准的内容局限于某一种技术,把中文新闻信息内容的概念模式和新闻元数据独立出来,并将新闻信息内容的 XML 格式和 JSON 格式单独进行规范和定义,从而使标准的范围更加明确,定义更加清晰。

GB/T 20092 旨在定义普遍适用于中文新闻信息内容采集、编辑、生成、发布、交换、存储、检索、评估反馈等新闻信息生命周期的各个环节的描述规范,为适应最新技术在新闻媒体中的实际需要,使标准更具有可用性,将标准名称修改为《中文新闻信息内容》,拟由四个部分构成。

- 第 1 部分:概念模型。目的在于建立新闻行业中文新闻信息的抽象概念模型,规范中文新闻信息文档的总体结构和描述方式,为新闻内容使用者提供指导。
- 第 2 部分:新闻元数据。目的在于规范描述中文新闻信息内容的各类元数据,便于在新闻媒体行业信息系统中实现信息内容的统一表达,一致描述。
- 第 3 部分:数据交换的 XML 格式。目的在于规定基于 XML 的中文新闻信息数据交换格式,为新闻媒体信息系统的建设提供技术支持。
- 第 4 部分:数据交换的 JSON 格式。目的在于规定基于 JSON 的中文新闻信息数据交换格式,为新闻媒体信息系统的建设提供技术支持。

中文新闻信息内容

第2部分：新闻元数据

1 范围

本文件规定了中文新闻信息内容中新闻元数据的基础公共元数据结构、文档层元数据、稿件层元数据和内容项层元数据。

本文件适用于新闻信息内容的采集、编辑、生成、发布、交换、存储、检索、评估反馈等新闻信息生命周期的各个环节。也适用于报刊、广播、电视、通讯社、新闻网站、社交媒体等新闻内容提供商及媒体应用与研究机构。

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 4880.2 语种名称代码 第2部分：3字母代码

GB/T 4881 中国语种代码

GB/T 20092.1—2022 中文新闻信息内容 第1部分：概念模型

GB/T 20093 中文新闻信息分类与代码

3 术语和定义

GB/T 20092.1—2022 界定的术语和定义适用于本文件。

4 基础公共元数据结构

4.1 地点

4.1.1 地点元数据结构

地点元数据结构定义见表1。