

ICS 35.240.30
L 67



中华人民共和国国家标准

GB/T 20092—2013
代替 GB/T 20092—2006

中文新闻信息置标语言

Chinese news markup language

2013-12-31 发布

2014-07-15 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
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 规则和约定	2
5.1 拼写规则	2
5.2 元素表达及描述要求	2
5.3 XML 组件的命名	3
5.4 CNML XML Schema 的版本和变更约定	3
6 CNML 体系结构	3
6.1 原则	3
6.2 总体结构	3
6.3 文档层结构	4
6.4 稿件层结构	5
6.5 内容项层结构	7
6.6 元数据结构	7
6.7 受控词表结构	9
6.8 关系结构	9
7 CNML 公共结构和基本数据类型	10
7.1 目标命名空间声明	10
7.2 基本简单类型	10
7.3 公共结构类型	14
7.4 复杂元数据项类型	32
7.5 元数据相关的 Schema 定义	82
8 CNML 的关系机制	91
8.1 关系概述	91
8.2 关系定义	91
9 受控词描述机制	98
9.1 概述和分析	98
9.2 受控词表的建模	99
9.3 受控词引用的建模	114
9.4 受控词表的例子	116
10 数字签名	119
10.1 概述	119

10.2	元数据描述	119
10.3	数字签名样例	122
11	CNML 文档层结构	127
11.1	概述	127
11.2	CNML 元素(〈CNML〉)	128
11.3	词表编目结构	131
11.4	信封结构	133
11.5	稿件容器	172
11.6	稿件关系容器	191
12	CNML 稿件层	194
12.1	稿件层定义	194
12.2	标准元数据组	202
13	CNML 内容项层	344
13.1	CNML 内容项层结构	344
13.2	内容项层实体内容	349
13.3	内容项层的元数据	363
14	CNML 内联智能标记	462
14.1	概述	462
14.2	内联智能标记的命名空间	462
14.3	规范说明	462
14.4	内联智能标记类型(cim: IntelligentMarkupType)	462
附录 A (资料性附录)	CNML 应用实例	474
附录 B (资料性附录)	CNML 标准元素中英文对照表	497
附录 C (资料性附录)	CNML 类继承总图	503
附录 D (资料性附录)	受控词表	504
D.1	受控词表的完善和更新	504
D.2	CNML XML Schema 标准受控词表	504
附录 E (规范性附录)	CNML XML Schema	518
附录 F (规范性附录)	内联智能标记 XML Schema	597
参考文献		600

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准代替 GB/T 20092—2006《中文新闻信息置标语言》。

本标准与 GB/T 20092—2006 相比的主要变化如下：

- 删除了引言；
- 修改了第 1 章范围里的部分内容；
- 修改了第 2 章规范性引用文件的引导语，删除了本标准中未直接引用的文件；
- 修改了第 3 章的部分术语定义；
- 修改了第 6 章 CNML 体系结构中的图 1、图 2、图 4、图 7，以及附录 C 中的 CNML 类继承总图；
- 修改了 6.4，删除了原稿件结构中与管理有关的内容，并删除了 12.1.2；
- 删除了 7.3，简化了关系的表达机制，并修改了第 8 章；
- 修改了第 9 章，删除了受控词的关系机制，修改了受控词的基类型，增加了对受控词表简化使用的机制和受控词语义描述的功能；
- 增加了 11.3，词表编目结构；
- 修改了 11.4，修改了标准的信封结构的部分定义以及有关应用示例；
- 修改了 11.5.2.3，修改了引用稿件类型定义；
- 增加 12.2.4.4，定义了稿件专题信息的相关类型；
- 修改了 12.2.5.8，增加了权利元数据组中授权使用信息下的元数据；
- 增加了 13.2.2.8，增加了对内容项分组引用和描述的结构；
- 修改了 13.3.4.4、13.3.4.5，增加了音频、视频元数据组中的元数据项；
- 增加了 13.3.4.6，新增了音视频合一的元数据组；
- 删除了附录 A.2，修改附录 A.1 内容并作为新的附录 A；
- 修改了附录 B，增加了部分元素；
- 在附录 D 中增加了关系词表等受控词表；
- 修改原有的附录 F，改为定义内联智能标记的 XML Schema；
- 针对以上修改和标准应用中提出的其他合理修改建议，对标准 Schema 作了相应修改，并根据由此带来的变化，对相应的图表和内容进行了修改。

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：新华通讯社、清华大学、北大方正电子有限公司、中科大洋科技发展股份有限公司。

本标准主要起草人：武国卫、刁毅刚、黄菁、李涓子、张鹏、曹学军、王付生。

本标准所代替标准的历次版本发布情况为：

- GB/T 20092—2006。

中文新闻信息置标语言

1 范围

本标准规定了新闻行业中文新闻信息的体系结构、公共结构和基本数据类型、关系机制以及受控词描述机制,并给出了新闻信息的数字签名和内联智能标记的使用方法。

本标准适用于新闻信息电子数据的“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。

本标准的使用对象包括报刊、广播、电视、通讯社、新闻网站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 4880.2—2000 语种名称代码 第2部分:3字母代码

GB/T 4881—1985 中国语种代码

GB/T 7408—2005 数据元和交换格式 信息交换 日期和时间表示法(ISO 8601:2000, IDT)

GB/T 18793—2002 信息技术 可扩展置标语言(XML)1.0

GB/T 20093—2006 中文新闻信息分类与代码

3 术语和定义

GB/T 7408—2005 和 GB/T 18793—2002 中界定的以及下列术语和定义适用于本文件。

3.1

元元数据 meta metadata

定义和描述元数据的数据。

3.2

受控词表 controlled vocabulary

从专业的角度将同一范畴的主题进行归类后形成的词表,并允许以受控的方式进行词条修改和扩展。

3.3

CNML 文档 CNML document

符合本标准的 XML 实例文档。

3.4

稿件 item

文档中包含的可管理业务对象,它一般有完整的元数据和内容,表示一个完整的新闻信息或其他定义信息。