

ICS 35.240.20
L 76



中华人民共和国国家标准

GB/T 20916—2007

中文办公软件文档格式规范

Specification for the Chinese office file format

2007-04-30 发布

2007-09-01 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	I
引言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 整体结构	2
5 文档格式标准 schema 格式	3
5.1 文档格式标准 schema 基本结构	3
5.2 文档格式规范元素和类型说明	7
6 文档存储格式规范	259
6.1 文档存储规范基本结构	259
6.2 压缩和打包的实现方式	262
6.3 支持的文件格式	262
6.4 支持的压缩算法	262
附录 A (规范性附录) XML 文档类型定义	263
附录 B (规范性附录) 用户 XML 数据支持方案	500
附录 C (规范性附录) 多语种支持方案	502
附录 D (规范性附录) 功能扩展支持方案	504
附录 E (规范性附录) 域代码定义	506
附录 F (规范性附录) 数字格式的格式码描述	512
附录 G (规范性附录) 数字格式枚举说明	519
附录 H (规范性附录) 图案(纹理)填充枚举	520
附录 I (规范性附录) 预定义图形类别枚举	522
附录 J (规范性附录) 电子表格函数定义	527
附录 K (资料性附录) 可能涉及的专利清单	553
参考文献	554

前 言

本标准的附录 A、附录 B、附录 C、附录 D、附录 E、附录 F、附录 G、附录 H、附录 I、附录 J 为规范性附录，附录 K 为资料性附录。

本标准由中华人民共和国信息产业部提出。

本标准由中国电子技术标准化研究所归口。

本标准的起草单位：中国电子技术标准化研究所、中国科学院软件研究所、北京信息工程学院、珠海金山软件股份公司、无锡永中科技有限公司、北京红旗中文贰仟软件技术有限公司、北京九州软件有限公司、上海中标软件有限公司、国家应用软件产品质量监督检验中心。

本标准的主要起草人：吴志刚、李宁、丁蔚、陈建明、黄芳、王长胜、李健萍、成修治、赵菁华、左家平。

引 言

本标准的发布机构提请注意,使用者声明符合本标准时,可能使用涉及第 5 章、第 6 章、附录 A、附录 B 和附录 C 中有关内容的相关授权的和正在申请的专利。

本标准的发布机构对于专利的范围、有效性和验证资料不提出任何看法。

专利持有人已向本标准的发布机构保证,愿意同任何申请人在合理和无歧视的条款和条件下,就使用授权许可进行协商。该专利持有人的声明已向本标准的发布机构提交。

以下列出有关专利权利人的信息:

专利持有人:中国电子技术标准化研究所

联系人:赵菁华

通讯地址:北京 1101 信箱

邮政编码:100007

电子邮件:zhaojh@cesi.ac.cn

电 话:+10-84029795 13910006825

传 真:+10-84029948

网 址:<http://www.egs.org.cn>

专利持有人:北京信息工程学院

联系人:李宁

通讯地址:北京北四环中路 35 号

邮政编码:100101

电子邮件:ningli@public2.bta.net.cn

电 话:+10-64884878 13501218765

传 真:+10-64879080

专利持有人:北京九州软件有限公司

联系人:顾国泰

通讯地址:北京朝阳区北土城西路 3 号微电子大厦 B 座六层

邮政编码:100029

电子邮件:guotai@soft99.com.cn

电 话:+10-82993166

传 真:+10-82993297

网 址:<http://www.soft99.com.cn>

请注意除上述已经识别出的专利外,本标准的某些内容有可能涉及专利。本标准的发布机构不承担识别这些专利的责任。

中文办公软件文档格式规范

1 范围

本标准规定了基于 XML 的中文办公软件文档格式标准。此处的中文办公软件包括文字处理、电子表格和演示文档应用。

本标准适用于中文办公软件文档格式和存储格式的定义,支持不同的中文办公软件之间的兼容和文档互换。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本标准,然而,鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本标准。

GB 13000.1—1993 信息技术 通用多八位编码字符集(UCS) 第一部分:体系结构与基本多文种平面(idt ISO/IEC 10646-1:1993)

GB/T 18792—2002 信息技术 文件描述和处理语言 超文本置标语言(HTML)(idt ISO/IEC 15445:2000)

GB/T 18793—2002 信息技术 可扩展置标语言(XML)1.0(W3C RFC-xml;1998,NEQ)

3 术语和定义

下列术语和定义适用于本标准。

3.1

可扩展置标语言 eXtensible Markup Language; XML

定义语义标记的规则,这些标记将文档分成许多部件并对部件加以标识。作为元标记语言,用于定义与特定领域有关的结构化标记语言的句法语言。

3.2

命名空间 namespace

在 XML 文档中,对元素和属性通过赋予一个有两个部分名称进行标识的手段。第一部分是名字空间,第二部分是功能名称。名字空间标识一组名称,以防止出自不同来源去纳入同一 XML 文档中的具有相同功能名称的多个对象发生混淆。各种典型的命名空间引用统一资源定位符(URI),因为每一个 URI 都是惟一的。参见参考文献[4]。

3.3

文档类型定义 schema

一种用于限定文档结构(如元素的顺序、出现次数、属性等)的机制,用于描述一类实例文档的结构。分析器可以根据 schema 来验证文档。本标准中特指 W3C 制定的 Schema 标准 REC-xmlschema-0-20010502 第 0~2 部分。

3.4

元素 element

在诸如 HTML 和 SGML 之类的标记语言中,一组标记、标记之间的内容以及标记所含的所有属性的组合。