

2006年中文办公软件文档格式XML标准研究六 PDF转换可能丢失图片或格式，建议阅读原文

[https://www.100test.com/kao\\_ti2020/64/2021\\_2022\\_2006\\_E5\\_B9\\_B4\\_E4\\_B8\\_AD\\_c40\\_64750.htm](https://www.100test.com/kao_ti2020/64/2021_2022_2006_E5_B9_B4_E4_B8_AD_c40_64750.htm)

一、问题的提出 办公软件是日常办公常用的软件工具集合，目前标准办公软件的内容包括文字处理、电子表格和演示文档。在20世纪90年代，微软凭借其在操作系统上的垄断地位，同时建立了在办公软件上的权威，使得其办公软件，包括文字处理软件Word、电子表格软件Excel、演示文档Power Point成为其赚钱的利器。作为办公软件鼻祖的Lotus公司，也无法与其匹敌。微软办公软件的垄断地位隐藏了文档格式兼容的必要性，如果所有的用户都使用微软产品，那么微软的文件格式就成为一种事实上的工业标准。因此，微软对其文档格式一直没有公开，希望借助这种方式持续保持其竞争优势。中国国内的中文办公软件起源于DOS时代的文字处理软件，如WPS等。这些软件一度成为中文办公处理必备的软件。但是一方面随着DOS平台向Windows平台的迁移，微软在系统核心上加强了垄断；另一方面，微软推出了Windows平台上功能强大的办公套件，提高了中文处理能力，中国国内的自主知识产权的办公软件日渐式微，无法与强大的微软抗衡。这种现状的最终后果是国内广大的办公软件市场都将被国外知识产权的软件占领，不仅是产业发展受到影响，而且在办公软件资料保管、文档安全存储上也无法得到保证。但是，仅仅从办公软件本身上与微软竞争，基本上是不现实的，因为在操作系统核心上受制于人，应用软件本身的发展和可靠性都要受到影响。在这种形势下，国家推动了从操作系统到应用软件的整体规划。在操

作系统上，推动基于Linux核心的、具有独立知识产权的操作系统；在应用软件上，提出了支持包括办公软件、浏览器、个人信息助理等软件产品。特别针对办公软件，国家在电子政务采购中积极支持国产软件，在2002年的政府采购中，国产软件包括金山、中文2000等公司拿到了超过一半份额的订单，而且随着永中Office的逐渐成熟，电子政务中使用的国产中文办公软件将会进一步增加比重。办公软件的多样化带来的直接问题是格式的不兼容。在电子政务采购的过程中，如果同时采购了几家国产办公软件产品，则在公文交换的时候，不同的软件之间无法打开对方的文件。解决这个问题有两个方法，一种方法是电子政务应用指定选择唯一的软件工具；另一种方法是在不同的软件工具基础上制定标准，要求不同的办公软件必须能够存储和阅读特定格式的文档。不难看出，第二种方法相对可行。因为如果指定唯一的办公软件作为电子政务中的应用选择，无疑将公文文档过度依赖于一个公司的产品，这显然是不合适的。因此为办公软件格式建立一套标准提上了研究日程。根据目前信息技术和办公软件开发技术的发展趋势，选择采用XML语言、确定中文办公软件的功能和格式需求、建立适合办公文档格式的对象模型，定义文档格式的Schema描述来解决格式标准化问题。在这种条件下，文档如果符合格式定义Schema，则在软件之间就是可以共享、阅读和处理的。

## 二、标准描述语言XML文档格式标准的描述可能有多种方法，例如，可以通过描述文档的二进制结构来完成格式的定义。但是就发展趋势而言，采用可扩展置标语言XML（eXtended Markup Language，XML）语言作为办公文档格式描述的方法是一种最佳选择。

### 1．XML语言

的产生XML语言诞生于1998年，是早期著名排版描述语言SGML的子集和简化。SGML通过90年代初面向格式的描述语言HTML逐步转化成功能更加强大的面向内容的描述语言，这就是XML。XML以其简单性、高效率、优秀的可扩展性和互操作性等优点，辅以技术成熟、支持 ISO/IEC 10646 (Unicode) 国际标准等特点，成为当今国内外电子商务与电子政务的关键信息交换技术之一。XML语言可以用来作为保证中文办公软件兼容的文档格式标准描述工具。

## 2. XML语言规范集合

XML语言本身包含了一系列的规范集合，集合中的各种规范相互作用，共同构成文档格式标准化的描述工具。

100Test 下载频道开通，各类考试题目直接下载。详细请访问 [www.100test.com](http://www.100test.com)