

●刘瑛 黄奇 王荣井

语义 Web 技术及其对电子政务的影响

摘要 语义 Web 是以机器自动处理信息为目标的新一代 Web 技术,它为电子政务的实施提供了新的技术支撑。语义 Web 的关键技术是信息描述语言标准,目前有 XML、RDF 和 OWL。语义 Web 技术对电子政务的影响主要有:解决电子政务中标准化问题,提供智能的查询服务,提供一站式服务以及实现有效的知识管理。图 1。参考文献 10。

关键词 语义 Web 电子政务 本体 知识管理

分类号 G250.76

ABSTRACT Semantic Web is a new generation of Web technology with the objective of automatic information processing, and can provide new technological supports for the implementation of e-government. In this paper, the authors introduce key technologies of Semantic Web, such as XML, RDF and OWL, and summarize its impacts on e-government, such as standardization, intelligent search and knowledge management. 1 fig. 10 refs.

KEY WORDS Semantic web. e-Government. Ontology. Knowledge management.

CLASS NUMBER G250.76

电子政务最重要的特征就是利用信息技术特别是互联网,将政府的管理和服务职能转移到网络上完成,实现政府组织机构和工作流程的优化重组,向全社会提供高效优质、规范透明和全方位的管理与服务。互联网是电子政务得以实现的技术平台。然而,目前的互联网主要是使用 HTML 来编制文档的,数据的显示格式和表达格式糅合在一起,无法实现计算机自动化处理,在信息资源的整合、异构信息平台的集成等方面能力也有限,电子政务的发展受到极大限制。

近年来,许多研究人员开始考虑使用机器可处理语义来丰富互联网上的可用信息,提高互联网上计算机自动处理的能力,提出了下一代互联网——语义 Web 的概念。语义 Web 技术的发展对电子政务的发展将产生巨大影响。

1 语义 Web

语义 Web 的概念最早是由互联网的创始人 Berners Lee 提出的,与目前的 Web 相比,语义 Web 不仅是一个人类可读的网络,还是一个机器可理解的网络。Berners Lee 为未来的 Web 发展提出了基于语义的体系结构——语义 Web 体系结构^[1],如图 1 所示。

在该体系结构中,第一层采用统一资源标识符 (URI) 来标识资源和属性,采用统一编码作为字符的编码方案,是整个语义 Web 体系结构的基础,它使得

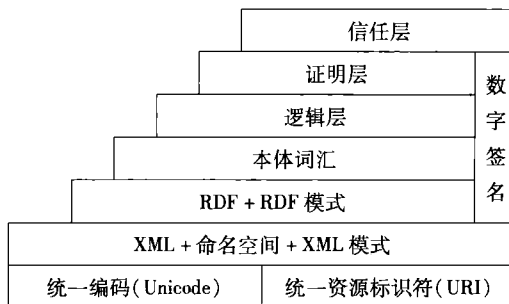


图 1 语义 Web 体系结构

信息的精确检索成为可能。第二、第三和第四层支持在语义上对网络信息进行描述,从而实现在此基础上更为复杂的逻辑推理操作,是语义 Web 体系结构的核心。第五层负责提供公理和推理原则,为智能服务提供基础。第六和第七层负责提供认证和信任机制。从第三层以上,所有对资源的描述都使用了数字签名技术,它是建立可信任网页的基础,语义 Web 的最终目的是要给人们提供一些可信任的应用。

2 语义 Web 关键技术

2.1 可扩展标记语言 (Extended Markup Language, XML)

XML 是 W3C 组织定义的计算机文档表示的标准标记元语言^[2-3],遵循清晰严格的语法,将数据的

显示样式与数据本身分离,具有良好的可扩展性。作为 Web 的知识表示语言,XML 允许使用者自定义标注。为了解决不同用户使用同样的标注名表达不同语义而引起的冲突,XML 提供了命名空间 (Name Space)。XML 的命名空间由 URI 索引确定,在 XML 文档中用作元素类型和属性名。通过在每个元素和属性前添加名字空间前缀,可以区别不同的元素集合。

为了让机器理解文档的逻辑结构并能自动地验证文档结构,W3C 开发了 XML 模式来描述数据的逻辑结构。XML 模式提供了一套完整的机制以定制文档的结构,它包含有两类基本约束:内容约束和数据类型约束。内容约束决定在哪里及什么时候可以使用元素;数据类型约束控制什么类型数据可以出现在元素中。

借助 XML,用户可以使他们的文件具有任意的结构,但却不能表示结构的意义。为此,W3C 提议用 RDF 描述信息的语义。

2.2 资源描述框架 (Resource Description Framework, RDF/RDFS)

RDF 是 W3C 推荐的一种信息描述方式^[4-5],目的是克服 XML 的语义限制,提供一种简单的模式来表示各种类型的数据。RDF 采用对象—属性—值三元组作为基本建模原语,每个三元组由相当于句子中的主语、动词和宾语 3 个部分组成。在 RDF 中,文档中的声明通常是某个事物 (人、网页或其他任何东西) 对于某些值 (另一个人、另一网页) 拥有某些属性 (例如“是老师”,“是作者”)。RDF 是一个开放的标准,用户可以根据应用需要,遵循 RDF 模式自行定义扩展模式。RDF 模式在 RDF 数据模型的基础上进一步定义了类和子类,属性和子属性,领域和范围等。

如果把 XML 看做一种标准化的元数据语法规范,RDF 就可以看做一种标准化的元数据语义描述规范。RDF 通过基于 XML 语法明确定义的结构化约定来帮助建立语义模式 (RDFS) 和语法编码 (XML) 之间的桥梁,以此来促进元数据的互操作能力^[6]。

2.3 本体 (Ontology) 和 OWL (Ontology Web Language)

XML 和 RDF 为描述语义网中的信息提供了元数据语法和语义上的规范。但是仅仅有元数据是不够的,网页上的元数据必须链接到特定的定义了元数据项和这些元数据项之间相互联系的文件上。这些共享的概念和概念间的互联的集合就叫做本体 (Ontology)。本体通过对概念的严格定义和概念与概念之间的关系来确定概念的精确含义,为某个领域

提供一个共享的通用的理解,是解决语义层次上 Web 信息共享和交换的基础。

本体通常用基于逻辑的语言来表示,可在类、属性和关系之间做出详细、准确、一致且完备的区别。OWL 是 W3C 最新推荐的描述语言的标准^[7],它在 RDFS 的基础上引入了大量描述逻辑的建模原语,增强了表述能力。OWL 有 3 个表达能力递增的子语言:OWL Lite, OWL DL 和 OWL Full,可以根据需要分别用于特定的实现者和用户团体。

3 语义 Web 技术对电子政务的影响

语义 Web 技术的发展为电子政务的发展带来了契机,各国政府纷纷关注它在电子政务中的应用。欧盟 eGOV 项目的 One-stop Government 的先导性计划以 GovML 来标识和描述政府公共化服务及日常生活事件;SmartGov 和 OntoGov 项目采用本体论语言来表示政府行政部门服务的本体并运用知识管理技术来串连行政服务流程以建立政府各部门服务执行机制的完整的运作平台^[8]。语义 Web 技术应用于电子政务将会对电子政务的发展产生以下影响。

3.1 解决电子政务中的标准化问题

电子政务处理过程中要求能够在异构平台、在不同的网络中实现数据交换和业务自动处理,必然涉及数据、公文和文档格式的标准化、统一化,需要建立一个能够描述政府部门内部、政府部门间和政府与公众间数据交换和业务处理流程的规范标准,以减少数据在处理过程中因标准不统一而引起的诸多问题。标准化是电子政务建设的基础性工作,是电子政务系统实现互联互通、信息共享、业务协同、安全可靠的前提。

作为语义 Web 关键技术之一的 XML 实现了数据内容和数据表现方式的分离,具有良好的自定义性和可扩展性,已经成为 Web 上的通用语言,是建立电子政务规范语言的基础。世界各国纷纷在 XML 标准的基础上,建立符合自己国家需求的电子政务规范语言。我国政府也已经着手研究基于 XML 的电子政务交换标准语言——cngXML^[9]。

在电子政务中采用基于 XML 的元数据标准,可以解决政府数据的统一接口问题,实现各级政府以及政府部门之间的数据共享和交换,避免数据资源重复建设。

3.2 提供智能的查询服务

电子政务管理模式下,用户可以通过政府网站提

供的查询服务来了解政府发布的各项信息。由于网络信息资源描述语言——HTML的局限性,现有的搜索引擎技术无法将不同网页上的相关信息整合在一起提供给用户,无法很好满足用户需求。

在基于语义的电子政务环境中,信息之间通过RDF实现了良好的语义关联,软件代理可以在本体的指引下查找准确信息,并能让本体语言和不同的信息关联起来,自动分析和整合信息,帮助用户只在某一个政府门户网站上直接获取合意的信息资源,而不必链接到别的网站。由于智能代理具有自动推理能力,还可以通过本体中的一般背景知识来推断一些更深层的知识,解答用户事先未能想到的一些问题,帮助用户更好理解政府政策。

基于语义的电子政务,要求政府各个部门在制定政策法规时不能互相矛盾,否则就会让公众对政府形象失去信心,损害政府形象。

3.3 提供“一站式”服务

通常情况下,用户要办一件事情,可能涉及政府的多个部门,需要多个政府职能部门的协同和配合,经过一系列的流程处理才会有结果。现在的网络环境中,用户还必须通过访问各个相关网站才能将事情办妥。例如,对于一个开餐馆的申请,申请者需要知道工商、税务、卫生、城管等各部门的审批和管理权限,随后按照传统的办事程序到各部门的网站递交申请,得到审批结果后再到另一网站重复同样的工作。这与电子政务的真正目的完全背道而驰,电子政务的真正目的应该是完全以用户需求为中心,向用户提供个性化的、一站式的政府服务。

语义Web通过建立基于本体的部门功能描述、分类体系,以及对政务处理过程的结构化、标准化和个性化处理,政府内部业务可以实现并行式办公,公众享受政府服务只需一个入口,不必按串行的工作方式去拜访,从而提高政府的工作效率。同样对于开餐馆的申请,在基于语义的电子政务环境中,申请者只需向电子政务门户网站提出要求,软件代理就会提示他必须具备的条件和需要提供的证明材料,然后在后端自动链接上述各部门办好各种手续,最后通知申请者全部手续已经办理完毕。这样,所有政府官僚体制被屏蔽,用户只需在某一点提出要求,并在同一点得到结果,真正实现政府的一站式服务。

3.4 实现有效的知识管理

电子政务实施过程中涉及政府、企业和公民的大量信息,实现有效的知识管理是电子政务成功的关

键。由于现有的互联网技术的局限性,政府的知识管理正面临着信息过量、无效的关键词检索、异构信息整合等问题。语义Web技术提供了解决这些问题的方法。

在语义Web中,通过XML、RDF和OWL在语法和语义上对信息资源进行描述,可以在此基础上利用本体编辑工具、本体存储和检索工具,实现对深度知识的自动抽取和维护^[10]。基于本体的信息导航和询问系统,可以快速而方便地找到所需要的知识,并且将最恰当的知识在最恰当的时间传递给最合适的行政人员,帮助政府实现最佳的行政决策。

参考文献

- 1 T Berners-Lee. Semantic web-XML2002. <http://www.w3.org/2000/Talks/1206-xml2k-tbl/Overview.html>
- 2 Tim Bray, Jean Paoli, C M Sperberg-McQueen. XML 1.0 recommendation. 2000. <http://www.w3.org/TR/2000/REC-xml-20001006>
- 3 D C Fallside. XML Schema part 0: primer. W3C Recommendation, 2001. <http://www.w3.org/TR/1999/REC-xml-names-19990114>
- 4 Frank Manola, Eric Miller. RDF primer. W3C Working Draft, 2002. <http://www.w3.org/TR/rdf-primer>
- 5 Dan Brickley, R V Guha. Resource Description Framework (RDF) model and syntax specification, W3C working Draft, 2002. <http://www.w3.org/TR/rdf-schema/>
- 6 朱礼军,陶兰,黄赤.语义万维网的概念、方法及应用.计算机工程与应用.2004,40(3)
- 7 W3C Recommendation. OWL Web Ontology Language Guide. <http://www.w3.org/TR/owl-guide/>
- 8 胡毓忠.电子化政府的新架构:创建以知识管理为主的政府信息网服务入口网站. <http://www.ntpu.edu.tw/pa/news/93news/attachment/931215/931215-4.ppt>
- 9 李磊,孙文兵.电子政务规范语言 cngXML. http://www.ccw.com.cn/htm/produ/topic/01_4_13_10.asp
- 10 D. Fensel, C. Bussler, Y. Ding, V. Kartseva, etc. Semantic Web Application Areas. <http://www.cs.vu.nl/mcaklein/papers/NLDB02.pdf>

刘 瑛 南京大学信息管理学系硕士研究生。通信地址:南京大学信息管理学系研二。邮编 210093。

黄 奇 国家信息资源管理南京研究基地主任。通信地址:南京大学树华楼313室。邮编 210093。

王荣井 国家信息资源管理南京研究基地研究员。通信地址同上。(来稿时间:2005-04-13)