

电子政务信息资源语义互操作模型研究 *

吴 鹏 高 升 甘利人

摘 要 信息资源语义互操作正成为语义网环境下电子政务新的研究课题。在对语义互操作模型研究与应用现状进行分析的基础上,基于元数据注册系统和元模型互操作框架,以政务异构信息资源的语义解析和整合为核心,提出电子政务信息资源语义互操作模型,设计语义解析和语义整合模型,并应用于国土管理信息资源的语义互操作。图6。表2。参考文献9。

关键词 电子政务信息资源 语义互操作 互操作模型 元模型互操作框架

分类号 G254

ABSTRACT To solve the problem of semantic understanding is the best function of semantic interoperability of e-government. It is the best way to solve the problem of semantic interoperability of e-government from the level of information resources. This paper analyzes the development of a model of semantic interoperability, proposes a model of semantic interoperability of e-government information resources based on the Metadata Registry (MDR) and Metamodel Framework for Interoperability (MFI), which take the semantic parsing and integration of heterogeneous information resources. 6 figs. 2 tabs. 9 refs.

KEY WORDS E-government information resources. Semantic interoperability. Metadata Registry. Metamodel Framework for Interoperability.

CLASS NUMBER G254

1 前言

当前,电子政务虽然已经实现了跨部门的信息交换,但是不同的信息系统建设模式和管理流程,使得不同部门的信息系统无法准确地“翻译”和“理解”信息的真实含义。在语义网络环境下,用户需要一个能够“理解”他们访问需求的电子政务系统,政府需要不同应用的信息系统能够“读懂”彼此信息内容。在目前开放的语义网络环境中,在电子政务存在大量异构和自主信息系统的现状下,面对不同的用户提出的个性化服务需求,如何解决这一问题?在努力消除“信息孤岛”的同时,却制造了一个个“语义孤岛”,这一问题已成为电子政务研究者研究

的热点。

电子政务专家 Sølvsberg 认为“信息语义反映信息之间的关系以及信息表达的含义。为了实现对于交互信息的相互理解,不同的电子政务系统必须共享一个代表信息含义的模型。语义互操作就是解决怎样达到这样互相理解问题的。”^[1]信息资源是电子政务的“食粮”与“血肉”,在电子政务环境下,整个电子政务系统完全是围绕着政务信息资源来发挥系统各项功能的,因此从信息资源层面解决电子政务的语义互操作问题是最佳途径。而基于数字图书馆互操作理论发展起来的元数据注册系统 (Metadata Registry, MDR)、扩展元数据注册系统 (eXtended Metadata Registry, XMDR) 和元模型互操作框架 (Metamodel Framework for Inter-

* 本文受国家社科基金“网站信息传递中符号表征的认知活动探索”(编号:08BTQ036)、南京理工大学科研发展基金“电子政务信息资源语义互操作模式研究”(编号:XKF09024)和南京理工大学经济管理学院预研基金“电子政务信息资源语义互操作研究”资助。

operability, MFI) 有效地支持了信息资源的语义描述、解析和整合。本文在此基础上, 基于韩国学者 Doo-Kwon Baik 等建立的元数据注册/元模型互操作协作模型^[1], 结合国土管理信息资源的实际情况, 分析电子政务信息资源存在的语义异构以及语义互操作应用需求, 提出电子政务信息资源语义互操作模型, 设计语义解析和语义整合模型, 并应用于国土管理信息资源的语义互操作。

2 电子政务信息资源语义互操作研究与应用现状

国外学者从电子政务信息资源语义互操作的定位、意义、管理模式到需要克服的技术问题等方面, 同时结合实际应用, 开展了较为深入的研究。其中具有代表性的研究成果总结见表 1, 相关典型研究项目见表 2。

表 1 关于电子政务信息资源语义互操作主要研究者和研究成果

研究者	主要观点和成果
Ralf Klischewski	指出元数据的语义互操作、基于本体的语义解析, 以及不同本体之间的互操作是实现政务信息资源语义互操作的关键因素(2003 年) ^[2] 。
欧盟电子政务互操作协会 (IDABC)	指出电子政务互操作包括组织、技术和语义三个层面, 认为语义互操作是与元数据相关的一种 Web 服务, 能够促进电子政务系统理解交换数据的涵义(2004 年) ^[3] 。
Peristeras	研究了电子政务语义互操作与组织、技术互操作的关系, 指出语义互操作是电子政务信息资源互操作的核心(2007 年) ^[4] 。
Luis Guijarro	政务信息资源语义互操作包括利用 XML 规范进行数据解释, 或者利用本体进行知识表示和开发(2007 年) ^[5] 。
Witold Abramowicz	指出语义互操作管理中关键因素是需要克服不同电子政务系统模型和框架的异构性 ^[6] 。

表 2 典型电子政务信息语义互操作项目

国家/组织	项 目
美国	电子政务数据参考模型 (DRM) 提供了一个在数据结构 (语法) 和含义 (语义) 方面达成一致的方法; Wordnet 语义本体英文语义解析提供自动化的方法, 解决了本体中词义关联问题。
欧盟	欧洲互操作性框架 (European Interoperability Framework, EIF), 这个标准为欧洲互操作框架范围内的语义互操作提供了明确的定义和执行计划。
英国	电子政务互操作框架 (E-Government Interoperability Framework, E-GIF), 包括: 电子政务的元数据标准 eGMS (eGov Metadata Standards), 主要解决句法描述问题; 集成公共部门术语 IPSV (Integrated Public Sector Vocabulary) 关于中央和地方政府行政活动的类属词典。
新加坡	ONTOGOV (Ontology-enabled e-Gov Service Configuration) 项目采用 OWLS (Web Ontology Language, 万维网本体语言) 和 WSMO (Web Service Modeling Ontologies, Web 服务建模本体), 并对 OWLS 的服务本体 (Service Ontology) 进行了扩展, 目标是建立一个基于知识本体的电子政务服务, 支持语义互操作。

但是, 无论是美国和欧盟的建设经验都没有提供一个统一的建设模型, 而是仅仅从语义

互操作的某个层面展开研究应用, 或是基于某种元数据方案之上的映射机制, 或是应用某种

语义描述语言来提高语义互操作性,或是从自然语义角度构建本体模型进行语义解析。在电子政务系统中,面对复杂的组织结构、灵活的控制结构和丰富的语义信息,我们需要解决什么样的语义冲突、消除什么样的语义异构以及如何将元数据和本体与信息系统有机集成,从而真正跨越组织和文化界限,实现语义互操作。回答这些问题就需要一个统一的信息资源语义互操作模型来支撑。

3 语义互操作模型的发展

在元数据收割和 OAI 框架体系的基础上,国外学者从人工语言角度进行语义互操作模型的研究,以注册体系为核心,相继提出了 MDR 模型、XMDR 模型以及结合软件建模技术 MOF (Meta-Object Facility) 和 MDR 模型的 MFI 等,它们之间的关系如图 1 所示。

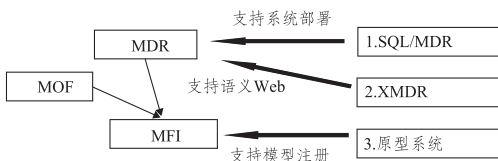


图 1 MDR、XMDR、MFI 相互关系^[1]

(1)MDR 模型与 XMDR 模型。MDR 通过对元数据定义、编码、转换、应用等规范进行发布、登记、管理和检索,以支持开放环境中元数据的发现、识别、调用及在此基础上的转换、挖掘和复用^[7]。由于 MDR 面向不同标准之间的数据互操作,在语义 Web 环境下,MDR 不能进行本体注册,不能描述属性关系的关联规则,同时不能支持 OWL 本体编码标准定义的规则和关系。美国劳伦斯伯克利国家实验室在 MDR 的基础上提出扩展的 MDR 模型,即 XMDR 模型^[8],它通过对现有元数据注册标准家族中的 ISO/IEC 11179 标准进行扩展,可以进行本体注册^[9],以及能够在本体的基础上构建 MDR;同时 XMDR 能和语义 Web 基于 XML 语言和 RDF 进行有机的集成,使得 XMDR 真正成为语义数据集成的核心。而且

XMDR 能够基于 UML 语言进行系统建模,从而使得基于 XMDR 原型构造的 MDR 系统能够在 MFI 体系下进行模型注册与互操作。

(2)MFI。XMDR 虽然扩展了 MDR 在本体注册方面的支持,但是缺乏对本体的互操作和演化管理,同时对过程的语义也没有管理模型。MFI 可以很好地解决这一问题,MFI 以网络复杂信息资源注册和资源(库)管理为核心,通过对不同类型的信息处理单元模型注册和管理,进行复杂信息资源语义互操作管理与服务。MFI 可以进行本体的互操作管理,为信息资源的语义互操作提供了概念模型,但是 MFI 模型仅从宏观方面进行了建模信息注册管理,在应用到实际系统建设时不能满足需要。而 XMDR 可以很好地解决这个问题,基于本体建立 MDR,支持 XML、RDF 等语义编码语言实现语义服务,同时 XMDR 模型可以注册到 MFI 模型,通过 MFI 模型完成本体的互操作和演化管理,这构成了政务信息资源语义互操作模型的两个核心研究内容,即基于 MFI 的电子政务信息资源语义注册模型和基于 XMDR 的电子政务信息资源语义服务模型。

4 电子政务信息资源语义互操作模型

电子政务信息资源语义互操作模型(图 2)以异构信息资源的语义解析和整合研究为核心。本体是语义概念解析的主要工具,语义解析研究以本体的应用和互操作为核心,通过元数据语义互操作、语境注释与互操作、本体互操作,实现“语义异构性的解析”。而概念解析最终的目标是为了支持语义异构信息的相互“理解”,通过基于本体的元数据注册系统的架构,以及元数据注册系统与描述语言的语义集成、与 Web 服务的语义整合,可以完成语法与结构互操作,实现“本地数据与远程数据的合一处理”。从资源管理的角度,基于扩展元数据注册模型、元模型互操作框架,采用模型驱动架构方法(Model Driven Architecture,MDA),结合联合检索模型和语义收割模型的优点,构建政务信息资源的语义解析与整合模型。

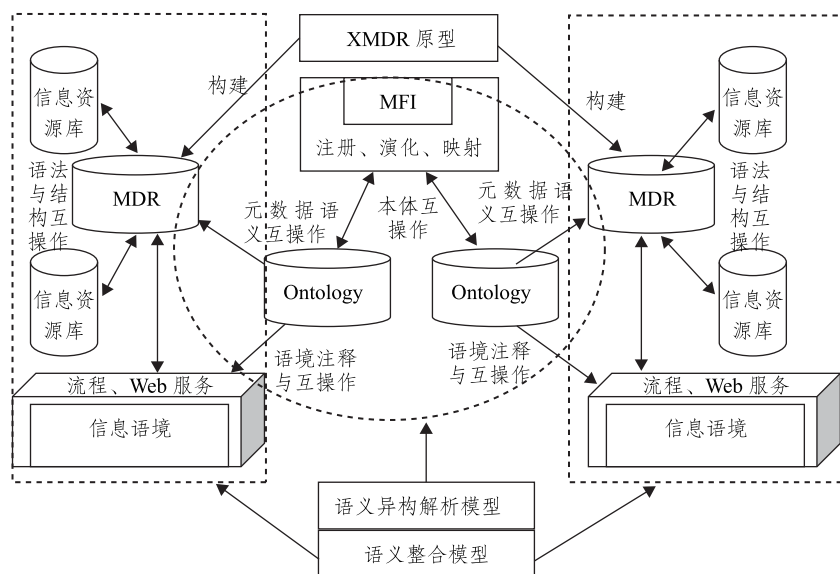


图2 电子政务信息资源语义互操作模型

4.1 电子政务信息资源语义异构分析

语义异构是服务于定制信息系统的自治数据建立和衍化的结果,元数据模型、元数据规范、对象的兼容性、数据格式是引起语义异构的主要原因,在电子政务语义互操作中,需要分析下列三类语义异构:

(1)语义相似的属性因为描述不同引起的语义异构,需要进行电子政务信息资源语义描述语言类型与功能分析,以总结语法和结构互操作需求。

(2)语义相似的实体因为描述模式不同引起的语义异构,需要进行电子政务元数据规范类型与功能分析,以总结元数据语义互操作需求。

(3)语义相似的实体因为提取解析模式不同引起的语义异构,需要进行电子政务本体模型的类型与功能分析,以总结本体互操作与演化需求。

4.2 电子政务信息资源语义异构解析模型

本体是语义概念解析的主要工具,以本体

的应用和互操作为核心,可以实现电子政务信息资源语义异构性的解析。对领域内同一概念体系的语义描述和解析,应用本体结合扩展元数据注册原型,构建电子政务元数据语义互操作模型;对于领域内不同概念体系的语义进行描述和解析,同时基于MFI实现本体互操作模型。具体包括:

(1)电子政务元数据语义互操作模型。根据电子政务元数据语义互操作功能的需求,我们结合联邦检索模型和元数据收割模型,以语义功能为中心,即以建立元数据注册系统、应用本体库为核心,以经过收割的分布元数据库提供主要的检索服务,但该元数据库可以是分布的,而且该体系并非要求收割所有资源站点的内容。对于内容或协议不符合要求的资源站点,用户可以通过查询服务代理,采用联邦搜索方式获取服务,而查询服务代理可以自动通过搜寻服务注册表进行判断(图3)。

(2)电子政务本体互操作模型。基于MFI,解决在电子政务系统中本体的注册、演化、映射及管理问题(图4)。

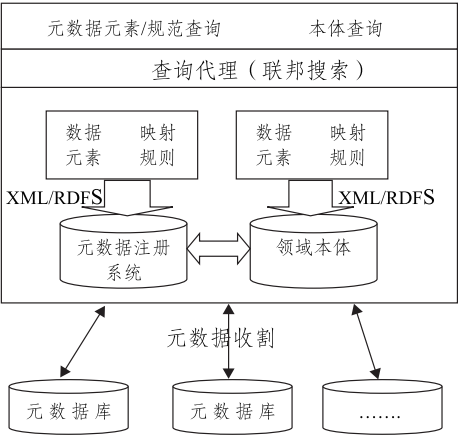


图3 电子政务元数据语义互操作模型

I 注册模式。基于 MFI 的核心模型(ISO/IEC19763-2),包括政务元模型的注册、目标、关系、模型分类器、层次匹配等功能对象规则,以及对象角色模型、属性转换模型。

II 演化模式。基于本体演化管理框架(ISO/IEC19763-3),包括本体注册模型、变化模型、约束模型、演化信息模型和变化传播模型,支持本体的基本内容注册、本体复用及演化。

III 映射模式。基于 MFI 的映射模型(ISO/IEC19763-3),包括电子政务领域映射规则、语义距离和相关性计算模式。

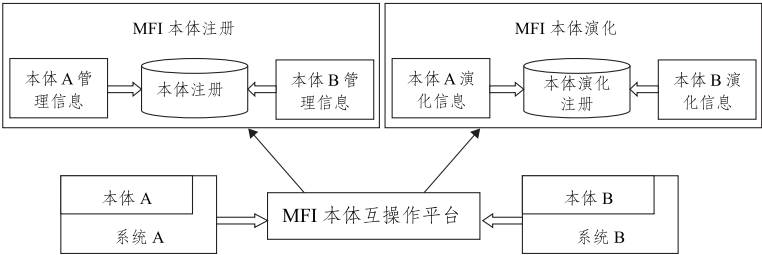


图4 电子政务本体互操作模型

4.3 电子政务信息资源的语义整合模型

信息是用户与电子政务系统以及不同电子政务系统的表现形式,概念解析的最终目标是为了支持语义异构的信息的相互“理解”,基于 XMDR 可以架构基于本体的 MDR,实现“本地数据与远程数据的合一处理”。

在语义整合模型(图5)中,本体定义了语义互操作的概念和关系,用来抽取用户查询中

的概念,同时和 MDR 中的数据元素信息进行整合。MDR 提供标准信息来整合异构数据库,并且管理异构数据库中的元数据数据元素的映射。结构映射主要将查询信息与数据库元数据元素进行映射,值映射主要解决概念相同的异构元数据数据元素之间的映射。具体内容包括三部分:

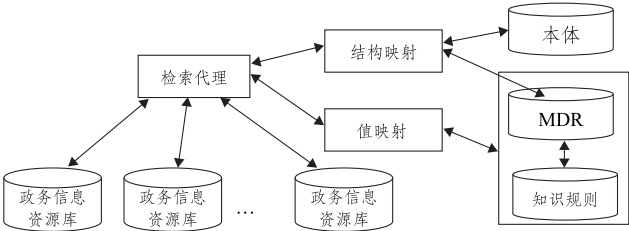


图5 电子政务信息资源的语义整合模型

(1)语义整合的构建模式和关联要素分析。包括本体和MDR在电子政务信息资源语义整合中的角色和作用分析,基于本体的MDR的构建模式,以及描述信息在MDR中的存储和查询模式。

(2)语义服务模式。解决电子政务信息资源语义描述和语义理解问题,研究本体、MDR与

语义描述语言的集成模式,基于本体的MDR与电子政务信息资源库的整合模式,解决语法与结构互操作问题。

(3)应用模式。以XMDR为核心进行元数据规范、概念关系、数据元素的内容解析转换,以UML和Protégé为工具进行元数据的语义描述,以Jena、lucene等本体工具进行语义解析。

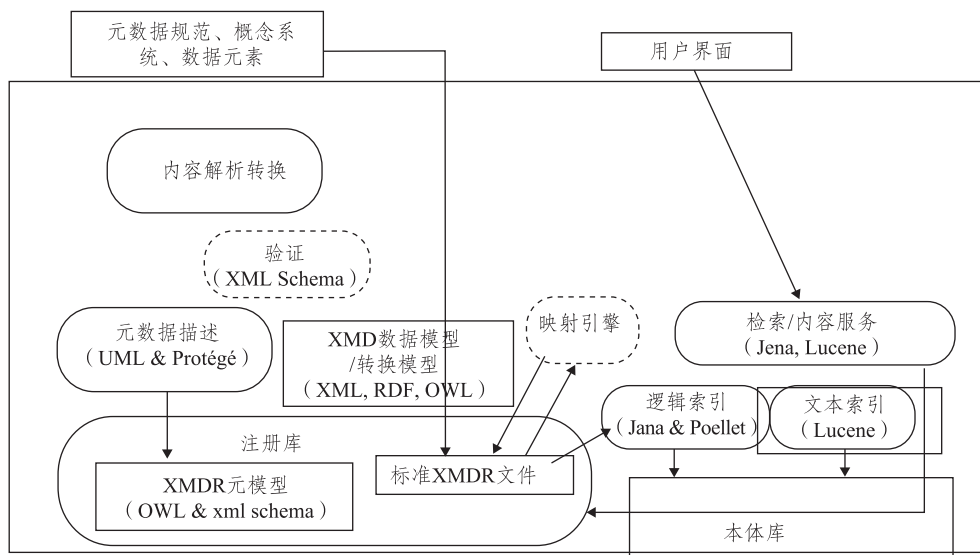


图6 国土信息资源语义整合应用模型

5 结语

欧洲电子政务互操作协会(IDABC)指出,语义互操作是电子政务信息资源互操作的核心^[3],电子政务系统不仅是信息设备的联合体,更可以看作是一个庞大的信息资源联合体。从抽象意义上看,这个联合体实际是一个巨大的符号系统,由无数的语法、语用和语义规则联结而成。语义和基于语义而实现的功能是其存在的目的和价值,所有的语法规则(各类协议、结构、约束等)和语用信息都是为了组织和传播语义而存在的。因此信息资源语义互操作正成为语义网环境下电子政务新的研究课题。

参考文献:

- [1] Doo-Kwon B, Jing Yixin, Kim Jangwon, et al. Metadata registries and the metamodel framework for interoperability[C]. Metadata Open Forum - "10 years of industry discussion". Sydney, NSW Australia, 2008(5).
- [2] Sølberg A, Hakkarainen S, Brasethvik T, et al. Concepts of enriching, understanding and retrieving the semantics on the Web[J]. ERCIM News, 2002(51): 41-42.
- [3] European commission. European interoperability framework for pan-European Government service[R], 2004.
- [4] Peristeras V, Liotas N, Tambouris E, et al. D2.7 study on interoperability at local and regional level[R]. Modinislot, 2007. (下转第122页)

《中华大典·哲学典》

《中华大典·哲学典》 任继愈主编 大16开 定价:8000.00元 2008年4月
由云南出版集团公司、云南教育出版社出版。

《中华大典·哲学典》是《中华大典》的一个子典,也是《中华大典》二十四个子典中第一个全部出齐的典。全书设三个分典:《儒家分典》、《诸子百家分典》、《佛道诸教分典》。每个分典下设典籍、人物、学派和范畴四个总部,近3200万字。

《儒家分典》主要收集儒家文献中的哲学资料。

《诸子百家分典》包括道、墨、名、法、杂、兵、天文、阴阳、农、医、纵横诸家文献中的哲学内容。

《佛道诸教分典》收集佛教、道教、中国基督教、中国伊斯兰教、犹太教、摩尼教、祆教等文献中的哲学资料。

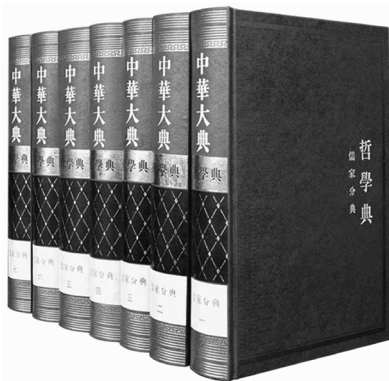
各分典之下的典籍总部收集有关主要哲学著作的序、跋、评介和著录等;人物总部收集重要哲学人物传记及有关个人生平的记载;学派总部收集有关该学派的综述性文字;范畴总部收集有关哲学范畴、命题或问题的文献。

联系方式:北京南天竹图书有限责任公司

联系电话:010-64130045, 010-64179066

地址:北京市朝阳区幸福二村首开幸福广场39号楼5单元5层

邮编:100028



(上接第82页)

- [5] Luis Guijarro. Interoperability frameworks and enterprise architectures in e-government initiatives in Europe and the United States[J]. Government Information Quarterly, 2007(1):89-101.
- [6] Abramowicz W, Bassara A, Wisniewski M, et al. Interoperability governance for e-Government. UNISCON 2008;14-24.
- [7] 张晓林, 梁娜. 元数据登记系统:基本概念与基本结构[J]. 现代图书情报技术, 2003(1):3-6.
- [8] Jeong Dongwon, Kim Young-Gab, Park Soo-Hyun, et al. Uniformly handling metadata registries[J]. Lecture Notes in Computer Science, 2006, 3647: 81-91.

- [9] Lee Jeong-Oog, Choi Yo-Han. Agent-based layered intelligent information integration system using ontology and metadata registry[C]// Intelligent Data Engineering and Automated Learning, Hong Kong, China, 2003:109-113.

吴鹏 南京理工大学经济管理学院讲师, 博士。通讯地址:南京。邮编210094。

高升 南京理工大学经济管理学院研究生。通讯地址同上。

甘利人 南京理工大学经济管理学院教授, 博士生导师。通讯地址同上。

(收稿日期:2009-08-30;

最后修回日期:2009-10-29)