

THE THEORY AND
IMPLEMENTATION
OF INFORMATION
INTEROPERATION SYSTEMS

信息互操作系统 理论与实现方法

贾利民 张遂征 袁宝军◎著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

国家“十一五”科技支撑计划课题（2009BAG13A01）

信息互操作系统 理论与实现方法

贾利民 张遂征 袁宝军 著

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书从互操作本体出发,以独立系统实现信息互操作为目标,介绍了互操作概念系统、互操作模型、互操作度量方法、互操作语言、互操作系统结构、互操作系统运行模式和应用案例,讨论了互操作路由、互操作集群、负载均衡、跨域互操作、安全互操作、互操作任务完整性控制、互操作安全控制等实现技术与方法。

本书内容新颖、结构完整,可供从事信息系统互操作研究、系统集成与信息整合的工程技术人员和相关专业师生阅读参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

信息互操作系统理论与实现方法/贾利民,张遂征,袁宝军著. —北京:电子工业出版社,2013.8

ISBN 978-7-121-20704-4

I. ①信… II. ①贾… ②张… ③袁… III. ①交通运输管理—管理信息系统—研究 IV. ①U495

中国版本图书馆CIP数据核字(2013)第130793号

责任编辑:赵娜

印 刷:三河市鑫金马印装有限公司

装 订:三河市鑫金马印装有限公司

出版发行:电子工业出版社

北京市海淀区万寿路173信箱 邮编 100036

开 本:720×1 000 1/16 印张:11.75 字数:263千字

印 次:2013年8月第1次印刷

定 价:42.00元



凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010) 88258888。

前言

随着信息技术在各个领域的广泛应用，信息系统一方面发挥了规范业务管理、优化业务流程、提高业务效率的作用，但由于组织管理与业务分工的关系，也形成了大量的“信息孤岛”和“业务孤岛”，信息系统效能未得到充分发挥。

互操作是实现信息系统之间业务协同、信息共享和信息交换的核心技术，是解决分布式、异构环境下系统整合难题的有效方法。为实现信息系统之间的互操作，国内外已经进行了大量的研究，但大多集中在数据交换和映射、语义理解与推理以及格式转换等技术实现方面。

2001 年，我们在组织中国铁路客票发售和预订系统升级改造时，为支持系统中的异地售票和售票存根汇集业务，采用了基于 Sybase Open Server 的交易中间件（Connection and Transaction Management Server, CTMS）和数据库传输中间件（Database Communication Server, DBCS）。当时，中国铁路客票发售和预订系统由 2000 多个车站售票系统、24 个地区中心系统、1 个部级中心和一个维护中心构成，全国部署有 10 000 多个售票窗口，是一个覆盖全国铁路的分布式客票交易系统，售异地票和隔夜统计清算是铁路客运提高服务水平的核心功能和提升管理能力的重要标志。CTMS 的部署结构如图 1 所示。

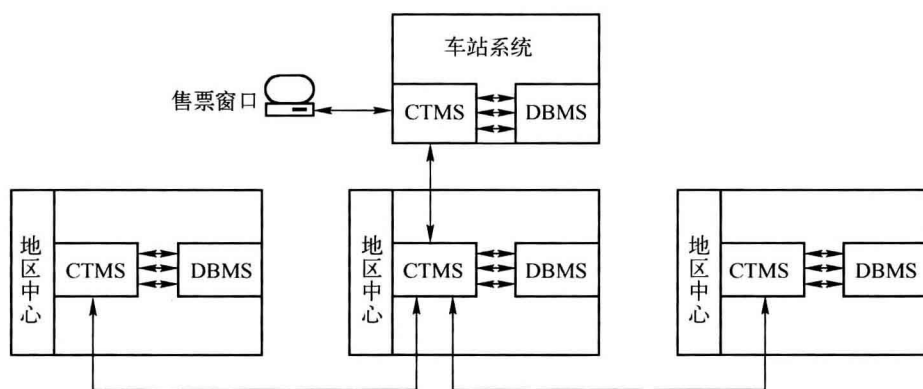


图 1 CTMS 的部署结构

在通过 CTMS 实现异地售票的同时，为保证售票交易的准确清算与审核，对异地售票的售票存根进行了多地存储（售票车站、售票窗口所在地区中心、票

源所属地区中心)。DBCS 部署在数据库服务器上,负责按统计时点(铁路行业的 18 点报)要求汇集所有售票存根。售票存根的汇集有四种情况:无中继单接收、无中继多接收、有中继单接收、有中继多接收,对应 DBCS 有三种运行状态:提取、中继、接收,如图 2 所示。



图 2 DBCS 的三种运行状态

运行在提取状态的 DBCS 负责提取指定票库中的售票存根,并按照定义发送给一个或多个接收存根数据的 DBCS;运行在中继状态的 DBCS 将接收到的存根数据按照定义转发给一个或多个接收存根数据的 DBCS;运行在接收状态的 DBCS 将接收到的存根数据存入售票存根库。具体到每个 DBCS 实体,可以根据定义运行在提取、中继、接收三种状态之一或这三种状态的组合状态。通过 CTMS 和 DBCS,实现了售异地票和售票存根多级汇总的功能,达到了各地售票系统业务协同的目的。

许多开发人员为了解决信息系统之间的业务协同问题,针对具体的数据交换和数据共享要求,与我们解决售异地票和售票存根多级汇总问题的方法类似,大多采用专用中间件方法来实现。在专用中间件应用实践的基础上,通过归纳和总结,产生了针对特殊类型业务的通用中间件,如交易中间件、传输中间件、消息中间件、工作流、数据库操作中间件等。这类中间件的产生,大大提高了业务协同实现的效率,简化了业务协同中间件开发的难度。

2009 年,我们在承担国家“十一五”科技支撑计划——国家道路交通安全科技行动计划“重特大道路交通事故综合预防与处置集成技术开发与示范应用”项目的“交通安全信息集成、分析及平台构建技术开发与示范应用(2009BAG13A01)”研究课题,构建跨部门交通安全信息集成与应用服务平台时,所面临的跨部门交通安全数据共享交换有其特殊的背景。

(1) 接入信息系统边界不明确。课题面对的是交通安全相关信息系统,但没有明确的具体对象,而且随着对交通安全问题的全面展开,涉及的将不只是交通运输和交通安全管理两个行业的信息系统,还将涉及铁路、民航、水运、机动车制造与维修、城建、保险、医疗、物流等众多行业的信息系统。

(2) 共享交换信息将不断扩展。课题需要交换的信息有初步定义,但不能

明确交换数据的具体形式、交换流程、控制要求和用户，随着交通安全研究对相关信息需求的深化和接入信息系统的增多，需要共享交换的数据种类和规模将不断扩展和完善。

(3) 信息系统支撑环境非常庞杂。共享交换数据的存储形式、应用支撑系统种类较多，几乎覆盖了目前市场上所有的操作系统、数据库管理系统和地理信息系统等。

(4) 信息安全防护系统各不相同。需要进行信息系统互联，但各行业采用的信息安全系统是不同的，不可能进行统一，但必须保证接入信息系统后原有的安全特性不被破坏、安全故障不被传递。

(5) 构建物理分设、逻辑集中的交通安全数据共享交换平台是课题研究目标之一，不需要构建数据中心或数据交换中心。

课题提出的交通安全信息集成与应用服务平台是一个复杂的信息系统互联系统，其基本拓扑结构如图 3 所示。

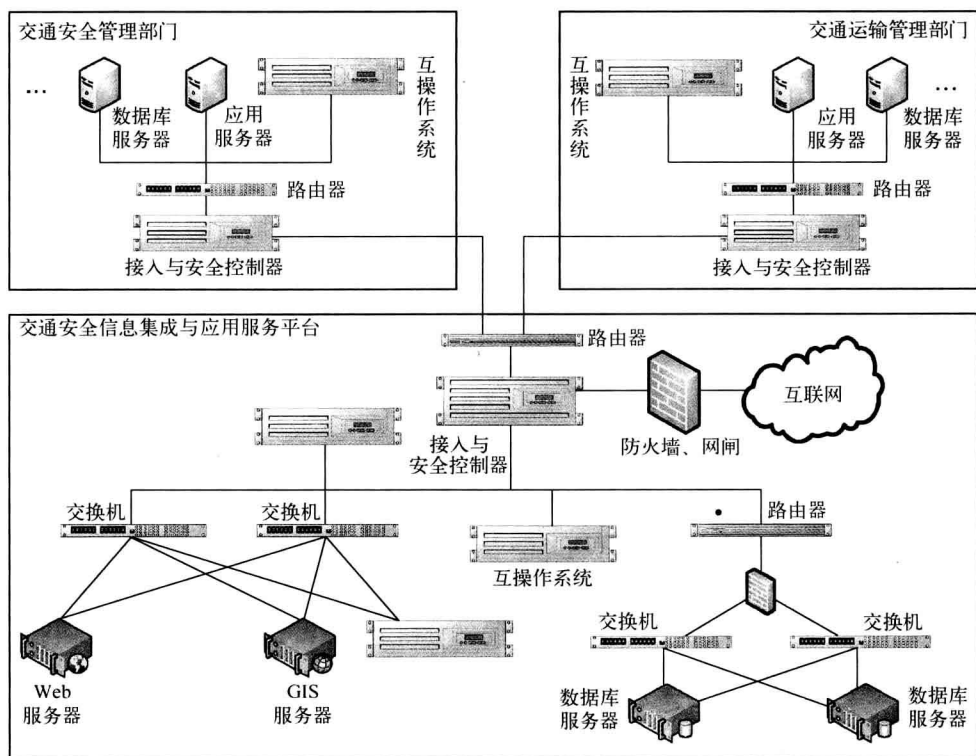


图 3 交通安全信息集成与应用服务平台拓扑结构

针对上述情况和我们在 863 课题“分布式交通系统信息互操作技术”中的研究进展,本书从互操作问题本身出发,研究和讨论互操作的一般实现问题。

一、互操作本体与互操作方法

以互操作本体为对象,以互操作的独立系统实现为目标,以互操作对象、互操作策略、互操作原语、互操作系统、互操作路由、互操作安全等为基本单元,构建互操作概念系统;基于互操作概念系统,从逻辑框架、本体结构、系统结构三个层面构建互操作模型;建立互操作定量度量模型和度量方法,通过互操作能力、互操作复杂性和互操作安全性等的定量度量,为互操作问题定量评估和改进提供手段和依据。

二、互操作实现方法

以互操作中数据抽取、转换、封装和导出四个基本环节所对应的四类基本操作的描述为基础,以互操作策略描述为目标,采用 SQL 样式建立互操作语言核心子集;针对互操作实现要求,讨论互操作路由、互操作集群、跨域互操作和安全互操作、互操作任务完整性控制、互操作安全性控制等核心技术的实现方法。

三、跨域互操作系统实现

以互操作的独立系统实现为目标,从技术与功能要求、系统结构、互操作流程、运行模式、安全管理、版本管理和任务管理等方面,介绍跨域互操作系统的基本构成和实现情况。最后,在对互操作系统应用进行分析的基础上,列举了三个简单的应用案例。

本书是国家支撑计划“交通安全信息集成、分析及平台构建技术开发与示范应用”(2009BAG13A01)课题中互操作技术研究的总结,引用了课题副组长包勇强(公安部交通管理科学研究所)专题组、课题副组长张可(交通运输部公路科学研究所)专题组、课题副组长王艳辉(北京交通大学)专题组、唐堃(北京宏德信智源信息技术有限公司)专题组的部分研究成果,还引用了吴旭任组长的 863 课题“网络化动态交通信息获取与交互技术”(2011AA110301)、“中国铁路客票发售与预售系统”、“城市运行保障和应急抢险车辆卫星定位管理物联网应用示范工程”的部分成果和资料,互操作系统原型系统主要由北京宏德信智源信息技术有限公司的葛大伟、南海斌、赵世雄、聂淼、贾献博、翟晓宇、贺丽红、丁

鹏、王建等负责开发，课题研究过程中，还得到了方守恩、王笑京、王长君、马林、关积珍、王云鹏、李宏刚等专家学者无私的意见和建议，这些对本书的完成提供了非常重要的帮助，一并致谢！

限于作者水平，不当之处在所难免，敬请读者批评指正！

张遂征

2013年4月9日于北京

目 录

第 1 章 概论	1
1.1 研究背景与意义	1
1.2 相关研究现状	2
1.2.1 本体 (Ontology) 与互操作理论	3
1.2.2 互操作框架研究	10
1.2.3 互操作模型研究	12
1.2.4 互操作技术与应用	14
1.2.5 互操作安全	16
1.2.6 互操作性度量	19
1.3 小结	22
第 2 章 信息系统互操作复杂性分析	23
2.1 交通管理信息系统	23
2.2 交通运输信息系统	26
2.3 铁路运输信息系统	30
2.4 信息系统互操作需求分析	33
2.5 互操作复杂性分析	36
2.6 小结	41
第 3 章 互操作本体与互操作方法	42
3.1 引言	42
3.2 互操作概念系统	42
3.2.1 互操作对象	42
3.2.2 互操作策略	45
3.2.3 互操作原语	48
3.2.4 互操作系统	49
3.2.5 互操作路由	51
3.2.6 互操作语言	55
3.2.7 互操作安全	59
3.2.8 互操作概念导引	61

3.3	互操作模型	63
3.3.1	互操作逻辑框架	63
3.3.2	互操作本体模型	70
3.3.3	互操作系统结构	71
3.4	互操作度量方法	78
3.4.1	互操作能力度量方法	79
3.4.2	互操作复杂性度量方法	80
3.4.3	互操作安全性度量方法	82
3.5	小结	84
第4章	互操作实现方法	85
4.1	引言	85
4.2	互操作语言实现方法	86
4.2.1	互操作语言元素	86
4.2.2	互操作语句跨域执行	90
4.2.3	互操作语句语法	94
4.2.4	互操作语句原型	98
4.3	互操作路由实现方法	100
4.4	互操作集群实现方法	103
4.5	跨域互操作实现方法	106
4.5.1	跨域互操作结构模型	106
4.5.2	跨域互操作功能模型	107
4.5.3	跨域互操作路由建立及维护	108
4.5.4	跨域互操作完整性控制方法	110
4.6	跨域安全互操作实现方法	113
4.6.1	跨域安全互操作结构模型	113
4.6.2	跨域安全互操作功能模型	115
4.6.3	跨域互操作安全性控制方法	115
4.7	小结	116
第5章	跨域互操作系统	117
5.1	概述	117
5.2	总体技术与功能要求	118
5.3	互操作系统物理结构	123

5.4 互操作系统总体框架	125
5.5 互操作系统总体功能	128
5.6 互操作系统总体流程设计	129
5.7 互操作系统运行模式	134
5.8 互操作安全管理	137
5.9 版本管理	139
5.10 任务调度优先级控制	140
5.11 小结	141
第 6 章 互操作系统应用示例	143
6.1 引言	143
6.2 应用关键技术分析	145
6.3 应用示例	148
6.3.1 道路交通全景信息平台	148
6.3.2 跨部门交通安全信息集成与应用服务平台	152
6.3.3 物联网车辆位置服务平台	158
6.4 小结	162
参考文献	163

第1章 概论

1.1 研究背景与意义

社会分工与协作是人类社会发展的基石，也是人类社会发展的永恒趋势。人类的社会分工提高了人类社会的生产效率，也产生了利益差别，并因利益冲突导致协作障碍。协作是使社会分工变得有效的基础，也是更进一步提高社会生产效率的必要条件，协作通过交换劳动成果使利益分配达到平衡。

随着信息技术的高速发展，以及信息技术不断渗透到人类社会活动的各个领域，为提高生产、管理和服务效率，各行各业已有大量信息系统投入应用。我国通过十多年的发展，特别是通过电子政务的“金”字工程和金融、交通、通信等领域的信息系统建设，取得了很好的社会经济效益。

伴随着大量信息系统投入应用，以及生产、管理与服务相关数据的大规模积累，因社会分工和组织管理的关系，信息交换已逐渐成为社会生产的关键环节，电子信息自动交换逐渐成为社会协作的重要形式和技术实现手段。

社会生产的各个方面是相互联系的，相互协作才能充分发挥信息系统的整体效能。为支撑不同业务和不同行业之间的信息交换，实现相关信息系统之间数据的同步、共享、交换和服务交互基础上的系统互联和业务协同，互操作理论与技术正得到越来越广泛的关注和重视。

近年来，物联网和云计算技术逐渐成为研究热点，各国从政府、研究机构到企业，都试图抢占新技术的制高点，成为产业革命的领头羊，带动国民经济可持续发展。从已经公开的研究成果来看，互操作要实现业务协同与物联网和云计算的目标是一致的，是物联网和云计算的基础技术，通过互操作，才能实现物联网和云计算所需要的并行计算、分布计算、网络存储、虚拟化和负载均衡。互操作也是物联网和云计算实现信息基础设施和业务应用开放的重要技术手段。

1.2 相关研究现状

互操作的本质是交换。《辞海》对交换的解释是“人们相互交换活动或劳动产品的过程。是社会再生产过程中连接生产和消费的一个环节。由生产决定，同时又反作用于生产、促进和阻碍生产的发展。只要有劳动分工，人们就必须相互交换活动……产品的交换是人们交换活动的一种形式”^[1]。应用于各行各业的信息系统是人类生产活动的一种基于效率的延伸，用于替代人类的部分生产活动。要使源自人类生产活动的信息系统充分发挥提高效率的作用，并使相关信息系统能够协调完成替代人的生产活动，就必须能够实现信息系统之间的信息交换，并使信息系统能够应用通过交换所获得的信息。

自有计算机系统以来，国内外就开始对互操作理论与技术进行研究。成为国际标准、得到大量应用、并在信息技术的许多方面发挥重要影响的当首推 ISO（International Organization for Standardization，国际标准化组织）的 OSI（Open System Interconnect Reference Model，开放系统互联参考模型）模型。OSI 最初是针对计算机网络互联需求于 1981 年制定的，分为七层，互联的两个系统每层的功能定义是相同的，如图 1.1 所示。

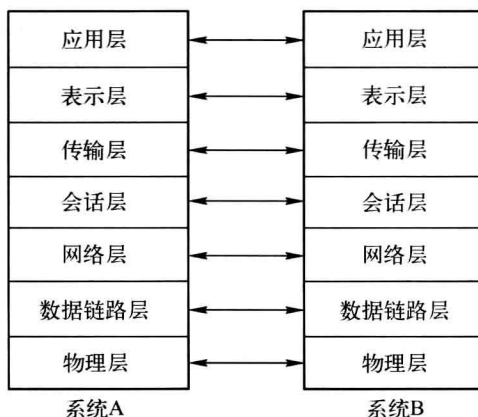


图 1.1 OSI 参考模型

OSI 的分层、两个系统同层功能相同、下层接收上层请求为上层提供服务的思想现已广泛应用于信息系统互联中的模型构造。

随着信息技术在人类生产活动中的大量应用，人类社会活动的各个方面迅速构建起了大量的信息系统。为解决信息系统之间的信息交换、协同运行等问

题,互操作理论与技术得到了广泛而深入的研究。国内外对信息系统互操作的研究主要从理论、框架、模型、技术和应用五个层面上展开。

1.2.1 本体 (Ontology) 与互操作理论

互操作理论的研究基本上是以本体论为基础,通过对人类社会活动客观存在的一般性把握,形成领域知识的概念体系(术语)和公理体系(关系),并在概念和公理基础上建立本体代数(规则),从而实现对互操作问题及其解决方法的一般形式化表示。

本体概念来自哲学,是只能用理性才能理解的本质^[1]。亚里士多德认为本体问题是关于本质、共相和个体事物的问题,是探讨本质与现象、共相与殊相、一般与个别等关系的第一哲学(百度辞典)。自本体论被引入信息科学领域以来,从本体定义、概念、关系、规则、语义、描述方法和针对具体领域本体的描述及其实际应用等各方面得到了广泛深入的研究^[3~39]。

哲学上本体论研究的目的是把握客观存在,这种对客观存在本质的把握不依赖任何语言和形式体系,不同的语言和形式体系对客观存在的本质把握应该是一致的,差异只在于表现形式。与哲学本体论不同,信息科学中本体论研究^[2]的目标是用形式化语言表述领域知识,是为了捕获相关领域知识、提供对相关领域知识的共同理解,确定该领域内共同认可的词汇(概念或术语——作者注),并从不同层次的形式化模式上给出这些词汇和词汇之间关系的明确定义。因此,本体论除研究建立领域知识和信息系统对象的一般概念、关系、规则、语言和形式化方法外,还研究具体领域知识的概念体系、分类系统、公理体系、本体代数和应用方法。

1. 本体构造

Gruber^[29]对信息科学中本体的定义是对特定领域中某套概念及其相互之间关系的形式化表达(Formal Representation),是对于共享概念体系的明确而又详细的形式化说明(维基百科)。对本体的描述通常从对象、类、属性、关系、函数、约束、规则、公理、事件等要素上展开,并采用相应的表示原语对本体各要素进行描述^[3,25]。

国内外研究者通过对特定领域本体的研究,已经建立了一些有重要意义的领域本体(维基百科)^[10,13,15,25],并采用图视方法对所研究的领域本体的概念、对象、类、属性、关系、函数、约束、规则、公理和事件等要素进行表示^[35],如图1.2所示的地球表面水及其模式的概念体系。

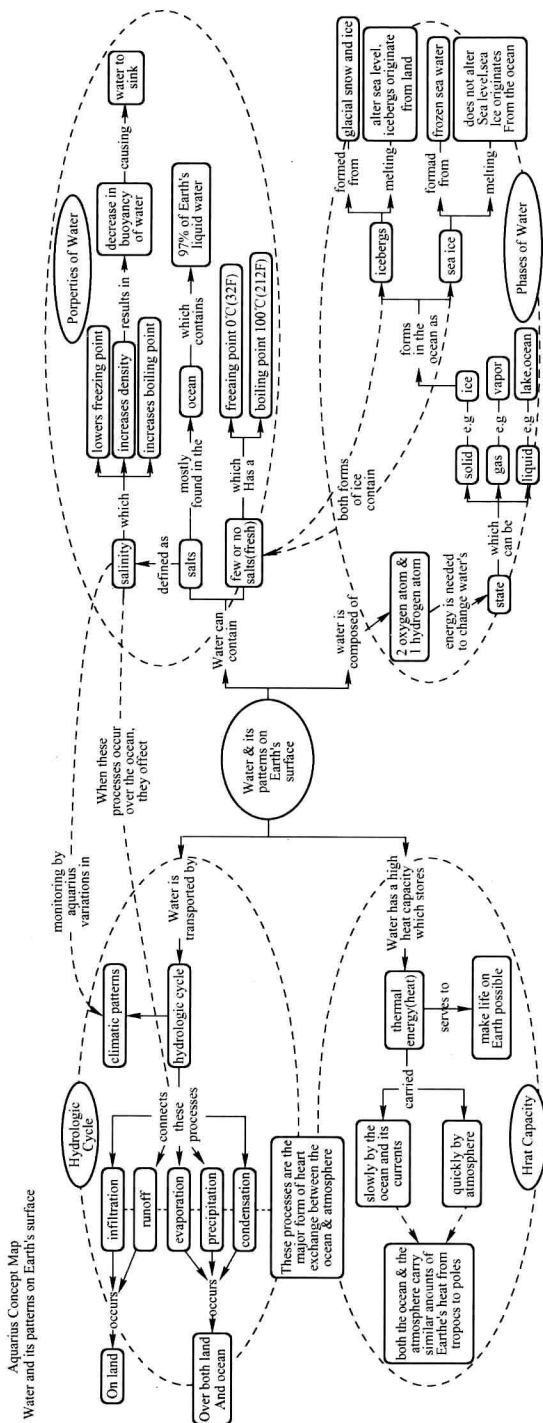


图 1.2 地表水及其模式

通过对本体及其形式化方法的研究,形式化描述已经取得比较丰富的成果,并在一些领域本体构建中得到应用。目前对本体的形式化描述还没有普遍接受的统一的方法和语言,国际上现阶段采用的形式化描述方法主要有: RDF (Resource Description Framework)、XML (eXtensible Markup Language)、SHOE (Simple HTML Ontology Extensions)、XOL (XML-based Ontology exchange Language)、OML (Ontology Markup Language)、CKML (Conceptual Knowledge Language)、OIL (Ontology Interface Layer)、DAML (DARPA Agent Markup Language) 和 OWL (Web Ontology Language) 等^[13]。

2004 年,国际 OMG (Object Management Group, 对象管理协会) 开始研究本体定义元模型 ODM (Ontology Definition Meta Model), 提出了以 UML (Unified Modeling Language) 为基础的本体承诺与表达语言体系 (Description Logic, DL; Web Ontology Language (W3C), OWL; Resource Description Framework (W3C), RDF; Simple Common Logic (ISO), SCL; Topic Map (ISO SC34), TM; Entity Relationship, ER) 及其各语言元模型之间映射规则,为本体描述开拓了广泛的可行途径^[3]。

本体通常由其基本成分构成,现在被普遍接受的本体形式化定义经过不断的完善,发展成为一个六元组^[2]:

$$\text{本体} = \{C, A^C, R, A^R, H, X\} \quad (1-1)$$

其中, C 为概念集; A^C 为概念属性集; R 为关系集, A^R 为关系属性集; H 为层次集; X 为公理集。

2. 互操作理论基础

信息科学研究本体的根本目的,是要利用本体论方法实现信息系统之间信息交换、服务协同和系统集成^[2],并使信息系统可以理解和应用所交换的信息、请求服务和获得服务结果、系统整合和业务协同。也就是说,信息交换、服务协同和系统集成的关键是信息系统的互操作问题。

信息系统之间的互操作问题的存在,主要是因为:

(1) 由信息系统替代执行的各项人类社会活动,本质上是相互关联的。相互关联的系统(人类社会活动)之间必然需要进行信息交换和业务协同。

(2) 信息系统的构建伴随着信息技术的发展和进步。不同时期建设的信息系统会因技术水平、认识能力、业务需求和组织管理等的差异和变化,导致信息系统对处理对象的理解、描述、记录、运用等方面的差异。

(3) 信息系统的构建与运行依赖于信息基础设施。针对业务覆盖范围、业务组织模式、业务处理时效、业务管理安全等特征和要求,以及在建设信息系统时设计人员的指导思想差异,将导致信息系统构建在不同的信息基础设施之上,并使信息基础设施以不同的方式互连和发挥作用。

(4) 相同类型的信息基础设施并不能完全兼容。例如,计算机、操作系统、数据库管理系统、中间件、服务支撑系统(如地理信息系统)等,不同厂商提供的产品并不能完全兼容,将导致信息系统之间不能采用简单直接的方法实现信息交换。

(5) 信息系统的构建与运行依赖于对跨地域覆盖的方式。跨地域分布的系统之间必须进行异地数据共享和异地服务调用。

(6) 信息系统的构建依赖建设者的能力和智慧。针对相同业务,不同的建设者会因知识背景、业务理解能力、趋势预见能力等存在差异,而导致信息系统对业务的定义、对象结构的描述、数据的组织与表示、功能的实现、设施的选择、开发语言的运用等方面存在差异。

(7) 人类活动的社会分工,形成信息孤岛是一种必然结果。信息孤岛的形成与是否构建信息系统没有直接关系,恰恰是因为信息系统应用对提升业务效率和质量的追求,使信息孤岛问题得以更明显地暴露出来,并为信息互联提供解决的技术和手段。

正如张东^[40]在“论元数据互操作的层次”一文中总结的,互操作问题的根源在于不同的信息系统和不同的实现人员对客观对象的认识差异和应用需要差异,即表现在语义差别、使用中的差别、描述方法的差别、词汇的差别、款目与集合、多版本、多语种、句法描述的差别等各个方面。

ISO/TC211 认为:若两个信息系统实体 X 和 Y 能相互操作,则 X 和 Y 对处理请求 R 的含义、处理过程和处理结果具有共同的定义和理解,如果 X 向 Y 提出处理请求 R , Y 能理解 X 的请求并对 R 做出正确响应,并将处理结果返回给 X ,且 X 能理解从 Y 处所获得的处理结果的含义。

研究表明,信息系统互操作的问题,是由人类活动的目的和对客观事物的认识的差异所产生的。因此,要实现基于本体的互操作,首先需要找到信息系统之间本体构造的共同部分,即两个本体实例具有共同的且可逆的投影(映射),并基于共同部分建立本体之间的联系,或可以在两个本体实例之间建立一种映射关系,该共同部分称为核心本体,如图 1.3 所示。

基于本体的互操作分为三个层次:数据层、语法层和语义层。数据层和语