



# A Guide to the Future of XML, Web Services, and Knowledge Management

Michael C. Doconta

[美] Leo J. Obrst 著

Kevin T. Smith

岳高峰 邢立强 孙文峰 等译

张 爱 顾迎建 审校

## 语义网

——XML、WEB服务和  
知识管理的未来



中国科学技术出版社  
CHINA SCIENCE AND TECHNOLOGY PRESS



# 语义网

——XML、Web 服务和知识管理的未来

[美] Michael C. Doconta Leo J. Obrst Kevin T. Smith 著

岳高峰 邢立强 孙文峰 等译

张 爱 顾迎建 审校

中国科学技术出版社

CHINA SCIENCE AND TECHNOLOGY PRESS

· 北 京 ·  
BEIJING

图书在版编目(CIP)数据

语义网——XML、Web 服务和知识管理的未来/[美]达康塔(Michael C. Doconta) 奥波斯特(Leo J. Obrst) 史密斯(Kevin T. Smith);岳高峰等译. —北京:中国科学技术出版社,2009.2

ISBN 978-7-5046-5375-8

I. 语… II. ①达…②奥…③史…④岳… III. 语义网络-研究 IV. TP18

中国版本图书馆 CIP 数据核字(2009)第 002156 号

自 2006 年 4 月起本社图书封面均贴有防伪标志,未贴防伪标志的为盗版图书。

本书中文版由 John Wiley & Sons International Rights, Inc. 授权中国科学技术出版社出版,未经出版许可不得以任何方式抄袭、复制或节录任何部分。



WILEY

John Wiley & Sons International Rights, Inc.

www.wiley.com

Original title: The Semantic Web - A Guide to the Future of XML,  
Web Services, and Knowledge Management

Copyright © 2003 by Michael C. Doconta, Leo J. Obrst, and Kevin T. Smith

All Rights Reserved.

This translation published under license.

版权所有 侵权必究

著作权合同登记号:01-2009-0203

中国科学技术出版社出版

北京市海淀区中关村南大街 16 号 邮政编码:100081

|      |         |      |     |
|------|---------|------|-----|
| 策划编辑 | 林 培 孙卫华 | 责任校对 | 孟华英 |
| 责任编辑 | 林 培 王 强 | 责任印制 | 安利平 |

发行部电话:010-62103210 编辑部电话:010-62103181

<http://www.kjpbbooks.com.cn>

科学普及出版社发行部发行

北京蓝空印刷厂印刷

\*

开本:787 毫米×1092 毫米 1/16 印张:11.875 字数:282 千字

2009 年 3 月第 1 版 2009 年 3 月第 1 次印刷

ISBN 978-7-5046-5375-8/TP·358

印数 1-3500 册 定价:23.80 元

(凡购买本社的图书,如有缺页、倒页、  
脱页者,本社发行部负责调换)

## 译者序

在开展数据与知识管理标准化研究过程中，我们要调研大量知识管理实现的技术及技术标准，发现这《语义网》非常符合我们的需要。经常拿来翻阅，其中提到的专家、项目、技术标准又可以按图索骥，找出相关的技术资料。由于要经常性的对其中的技术及标准进行讨论，索性就翻译了部分关键内容。后来，发现这个翻译工作很有意义，既能互相勉励，又可以相互研讨，分工协作，共同提高。最后，课题组决定翻译全部《语义网》，这样还可以供别人共同参阅。接下来的工作就是分工、翻译、研讨；然后，就是分别汇总。最后，制作图表，反复校对、整理，统一词汇，统一术语。

语义网是为实现人与机器之间的相互“理解”和“对话”的技术桥梁。自从有了计算机以来，人们不断研发新的技术、创造新的理论，来实现人机对话的过程。人们开发了各种“语言”，最初的机器语言，到汇编语言，到后来的 C 语言、Java 语言，通过专业人员进行编程，将人们的意图翻译成机器能执行的代码，完成人们交给的计算工作。随着信息技术、通信技术及知识工程技术的发展，机器理解人们思想的能力也在逐步提高。通过关键字、主题词表，人们在互联网上可以找到自己所需要的内容，但与此同时，大量不相干的 web 网页也会出现。另外，信息化逐步深入社会生活的各个方面，电子商务、电子政务、企业信息化的广泛而深入的推进，产生了海量的信息资源。如何有效地找到所需要的、有用的资源，如何让计算机更精准地理解人们的意图，以及如何能在网络上更好地展现内容和知识，已成为需要迫切解决的问题。

《语义网》中系统地介绍了各种相关技术。从 Tim Berners - Lee 提出的语义网基本思想出发，系统地介绍了 XML、Web 服务、分类体系和本体等技术，深入浅出，其中涉及的各种技术的基本原理；提供了丰富的图表、示例、图解、说明等，方便读者学习和理解。另外，《语义网》中涉及大量的专著文献、领域专家、技术标准、实验室、科研项目、辅助工具等，对于初学者来说非常有用。读者可以此为索引，找到该领域的大量背景资料，从而将语义网相关的技术串联起来。

《语义网》适用于要了解知识管理、语义网的初学者、非专业人员。对于大部分初学者或非本领域的人员来说，语义网、知识管理技术涉及计算机技术、人工智能、信息资源、软件工程、语言学等多个专业领域，让人感觉很抽象，不容易理解。而这些读者没有或者不需要花费太多精力来学习和研究技术细节，《语义网》就满足了这部分读者的要求。通过对多种技术的相互对比，解释了各项技术的优缺点，可以帮助读者在短时间内快速理解相关技术的基本原理。

语义网将在我国逐渐发展并成熟完善。语义网技术在国外已经比较成熟，研究和应用都比较广泛。随着我国信息化的逐步深入，语义网技术也会向互联网技术一样在社会生活的各个方面得以普及和应用。对于大部分读者来说，《语义网》都可以成为启蒙性的读物。

《语义网》由数据与知识管理标准化研究课题组统一翻译，岳高峰及邢立强通稿，张爱、顾迎建审校。本书的第 1、第 2、第 9 章由邢立强负责翻译，第 4 章、引言由孙文峰负

责翻译，第5章、前言由吕成功负责翻译，第6章由汤艳丽负责翻译，第7章由孙广芝负责翻译，第3、第8章、致谢由岳高峰负责翻译。

最后，要感谢中国标准化研究院的张爱书记和全国组织机构代码中心顾迎建主任，是他们不断的勉励和支持，使我们的翻译工作得以在繁忙的工作之余，在双休日或在晚上的业务时间之中，不断推进；在他们的直接指导下，开展了一系列的研讨和交流工作。感谢全国组织机构代码中心的郭秀婷，中国标准化研究院高新所的刘植婷、姜晓琳，对我们的工作提供了指导、建议和协助校对等工作。此外，还有中国人民大学的彭愚、北京航空航天大学的孙垚同学，进行了部分排版及校对工作。

希望《语义网》能对那些想了解、学习和研究语义网、知识管理技术的人们有所帮助，希望能够吸引更多的人加入到相关研讨中来。

译者

2008. 12

yuegf@cnis.gov.cn

www.cnis.gov.cn

## 对语义网的赞誉

目前正在发生的一场革命,就是基于内联网、外联网或者互联网的链接,使网络变得有含义、可理解,并能够为机器所处理。这就是所谓的语义网,它将使人们以知识为中心对“每个事物”进行思考。本书之所以独特,就在于它囊括了语义网涉及的所有技术,包括语义网的范围、XML 以及所有关键技术和协议、Web 服务及协议、资源描述框架(RDF)、分类体系、本体、语义网的商业案例以及企业的实现路线。所有的组织、业务、业务经理、开发人员和 IT 专业人士都应该仔细地研读关于知识应用和实现下一代应用系统/框架/运动的研究内容。该项研究是令人激动的,它所提出的应用系统/框架/运动将使所有人受益。

**Stephen Ibaraki**

iGen 知识方案公司 主席及首席架构师

语义网根植于对有关语境中词汇的理解。本书恰恰扮演了这样的角色,通过提供相关技术的关键定义和新兴技术的词汇,以助于人们理解语义网和相关技术。

**JP Morgenthal**

AG 软件公司 首席服务架构师

本书献给 Tim Berners - Lee, 他描述了语义网的远景, 以及致力于将此远景转化为现实的人们。Vannevar Bush 正在注视着它的前景, 并微笑以对。

# 引言

“毁灭我存在的根源是做着  
我知道计算机可以为我做的事情”  
——Dan Connolly, 《XML 革命》

让人感到无比懊恼的莫过于：自己已经解决了某个非常复杂的问题，结果现在却找不到记录着解决方案的文档或笔记。通常会拒绝重新考虑解决方案，因为你觉得自己已经解决了问题，并不想再浪费时间重复以前的工作。事实上，极端一点来说，找出从前的解决方案可能比你重新去做花费更多的时间。这就是我们信息管理工具与信息存储能力不同步导致的直接后果。

以个人电脑为例，价值 1000 美元、硬盘存储量为 60 ~ 80G 的个人电脑，文件的存储能力（假定 1 字节为 1 字符，明文，每页 3500 字）大概为保存 1700 万到 2000 万页的信息。大部分是无法以明文格式检索出来的专属的、二进制格式的文件。因此，对于我们的个人信息来说，主要的知识发现方法是以随机的、等级目录结构的方式产生。把这个例子应用于公司，我们能看到存储容量、信息格式和访问途径的多样性通过员工数量的增加而扩大十到百倍。

一般来说，很明显我们只能主动地管理我们所生产的所有信息中的一小部分。这样就会导致失去生产力，降低税收。事实上，将信息通过选择、增补、排序、关联和评注的方式转化为知识是一种积极的信息管理方式。本书的目的就是展示一条清晰的路线，也就是使用语义网技术改善机构中的知识管理；其次，我们检查语义网的技术构件包括 XML（可扩展置标语言），Web 服务和 RDF；最后，我们不只是向大家展示语义网将如何运行，而且提供一些例证和商业案例来展示如何应用这些技术使投资获得巨大的回报。

## 现在为什么应该读本书

事物之间的相互关联因为一种潜在的诱人目的而变成一种趋势，一些个体参与者试图达到这个目的却常常只能完成某部分。例如，电子设备集中的发展趋势是以降低设备成本和提高实用性为目的。软件要素的发展趋势是以软件的再利用为目的，也就是为了降低成本和提高市场销售的速度。个性化结构的发展趋势是以个体授权、成就感和降低成本为目的。语义网的发展趋势则是以数据的语义互用性为目标，使其能够自主应用，改良搜索工具和机械的推理。

精明的机构对于这些强大的趋势是不会忽视的。另外，如果这些趋势影响或改善了任务评定的应用，那么迅速掌握一些东西就很必要，这就是语义网案例。当前的语义网正在逐步出现在数以千计的像图书科技、国防、医药和经济等产业的实验计划中。另外，像 IBM、HP 和 Adobe 公司的技术领军人物都在使用语义网产品，而且更多的 IT 业公司也拥有内部的语义网搜索方案。总之，语义网的关键领域超出搜索阶段，发展到执行阶段。

语义网的多米诺骨牌效应已经开始显现：从 XML 到 Web 服务、到分类体系到本体再到推论。这并不是在描绘最近的风尚，相反，这是在知识表现方式上的多年研究与实验的顶点。现在这个推动力就是世界范围网络的成功。HTML、HTTP 和其他网络技术为成功地进行信息共享提供了一个强有力的先导。已经存在的网络不会消失；语义网技术将会把知识共享与发现融入现存网络中。

## 进入这一复杂主题的途径

本书探讨问题的场景是 CIO 和 CEO 们为企业所关心的技术视角。在此场景中，首先用清晰的术语解释概念并列举出具体的示例。然后，对技术作出判断，而不是在为技术做宣传。本书中的一些内容可以更好些，也有评判优缺点和令人生厌的地方。最后，本书奠定了技术方针的基石，并把其放在本书最后一章。

每个主题的模板都是针对各个领域的关键问题提供出简单明了的答案。另外，我们为所有出现在书中的重要概念提供了具体而又有说服力的示例。本书还提供了大量的例证性的图表来辅助概念的解释说明。最后，本书展示了一些作者引入的新概念，来补充解释对这些技术如何解决，为什么这样解决的理解。

## 本书的组织结构

本书分为 9 章，可以按顺序阅读，也可单独阅读。

**第一章 语义网是什么。**本章描述了实现数据的机器处理的语义网远景，以及我们如何到达这一远景。本章还描述了实现语义网的基本框架、为什么需要语义网以及其他章节中关键技术如何用于语义网。本章还介绍了一些诸如“智能化数据连续体”以及“组合实验”等概念。

**第二章 语义网的商业案例。**本章清晰地描述了一些具体的案例，来说明商业机构如何利用语义网来提高竞争力。具体来说，本章给出了决策支持、业务发展和知识管理方面的案例。本章的结束部分给出了语义网技术的现状。

**第三章 理解 XML 及其对企业的影响。**本章解释了 XML 为什么能够成功，什么是 XML，什么是 XML 模式，什么是命名空间，什么是文档对象模型（DOM），XML 如何影响了企业的信息技术。本章包括了仅仅 XML 元数据为什么不够用、数据可靠性的发展趋势等。最后，以语义层次的概念介绍结束了本章。对于还没有全面应用 XML 的企业来说，本章是必读的一章。

**第四章 理解 Web 服务。**本章阐述了当前 Web 服务的各个方面，并讨论了 Web 服务的未来发展方向。本章说明了如何发现、描述并访问 Web 服务，以及这些功能隐含的技术。本章还提出了部署 Web 服务的具体案例，并回答了“为什么使用 Web 服务”这一问题。最后，本章对高级 Web 服务的应用进行了详细描述，其中包括编排和安全。本章最后以支持网格的 Web 服务和支持语义的 Web 服务方面的讨论结束。

**第五章 理解资源描述框架。**本章描述了什么是 RDF、RDF 模型与语法的区别、其功能特点，为什么对 RDF 的采用没有像 XML 那样被迅速，为什么这一状况将发生变化。本章还介绍了这一技术的新案例，称为“非文本建模”。本章以使用 RDF 模式方面的讨论结束。本章强调了数据项之间关系的清楚建模的重要性。

**第六章 理解字母汤的其他部分。**本章涉及与 XML 相关的各种技术，解释了 XPATH、XSLT、XSLFO、XQuery、XPointer、XInclude、XML Base、XHTML、XForms 和 SVG。除了直接、清楚地解释这些技术的目的之外，本章还给出了示例，对每项技术的实用性和应用前景进行了评判。

**第七章 理解分类体系。**本章说明了什么是分类体系以及分类体系是如何应用的。本章使用说明性的示例给出了分类体系的理解，并显示了其与本体的不同之处。本章介绍了具有一定难度的概念，称为“本体谱系”。然后介绍了分类体系的一项常见应用，称为主题图和 XML 主题图。本章给出了对主题图和 RDF 的比较以及他们当前一些特点的讨论。

**第八章 理解本体。**本章的介绍非常详细，并采用渐进式、按照不同的模块的方式解释了本体是什么，怎么应用，如何获得语义互操作性。本章的开始先给出一个具体的业务案例，继而将本体的定义从几个不同的角度进行分析。然后，解释了本体的一些概念，如语法、结构、语义、语用、内涵和外延等。本章还给出了这些方面的具体示例，包括软件智能体对这些技术的使用。在解释主题词表与本体的区别时，引入了一个较有深度的概念“重要性三角”。本章然后介绍了知识表示和逻辑来详细描述本体背后用于机器推导的应用系统概念。本章包括对当前本体语言（如 DAML 和 OWL）的详细解释，并对本体在机构应用前景方面做了评判。

**第九章 形成公司通向语义网的路线。**本章给出了在公司中利用前面章节中探讨的语义网技术的详细的技术路线。本章将大多数公司的信息和知识管理现状与以知识为中心的组织进行比较，给出了创建技术路线的大环境。本章详细描述了以知识为中心组织的关键过程，包括（知识的）发现、生产、搜索和获取以及成果的应用（包括重复使用）。然后，给出了成为以知识为中心的机构的详细步骤。这些步骤包括使命定义、培训需求、技术实施、员工和计划。最后，本章以宣传性的号召进行了总结。

本书是当今新出现的数据革命方面的一个全面而深入的基础性和战略性的教程。每一章都对技术及其当前的状态进行详细、真实并且权威的评估，并提出如何在机构中使用这些技术方面的建议。适当时，还突出显示了使用这些技术的“格言”或原则。

## 本书的读者

本书为管理人员、技术领导和高级开发人员提供了战略指南。某些章节对任何对语义网感兴趣的人都有用；某些章节在涉及了基本的内容之后，比较深入的讨论了相关主题。然而，本书中的章节并没有对相关技术做专业而深入的介绍。

虽然本书应该从头到尾地渐进式阅读，然而书中的个别章节更适合于特定的人群。高级管理人员可能更喜欢阅读侧重于战略性理解、业务案例、语义网路线的章（第一、二、九章）。CIO 和技术主管们可能愿意阅读所有的章节，但可能会感觉路线更有用（第九章）。培训经理们可能更关注于语义网关键技术的章节，如 RDF（第五章）、分类体系（第七章）和本体（第八章），并据此来设置培训日程。感兴趣的高级开发人员和开发人员应该阅读并理解所有的技术章节（第三章到第八章）。

## 附属网站中的内容

附属网站的网址为：<http://www.wiley.com/compbooks/daonta>，包括如下内容：

源代码。本书中所有程序清单的源代码都可从一个压缩的文件夹中获得。勘误表。读者或作者发现的任何错误都被列出，同时给出改正后的文本。第八章的代码附录。由于第八章的一些程序清单很长，在书中进行了删减，而在网站中它们被完整地给出。

联系地址。作者的 E-mail 地址以及一些常见问题的回答。

## 致 谢

因为本书的主题是一个重要话题，包括合著者的丰富学识，以及提供关键性和战略性指导的写书思路，这些令我在写书时就已获益匪浅。与此同时，我在写另一本书 *More Java Pitfalls* 时遇到了一些困难。在那个困难时期，我非常感激来自妻子黎妮以及孩子们萨曼塔和格利高里的支持。我有幸与她相伴，她是一位令人着迷的母亲和妻子。在写本书时，我们搬过家，每个人都知道搬家有多困难。我也要感谢我亲家，巴迪和雪莉·贝尔登他们那里的支持。感谢鲍勃·艾里奥特、艾米莉·赫曼、布赖恩·斯奈普、迈克尔·弗雷德里克等来自威力出版社的员工，感谢他们在整个过程中的理解和支持。这个计划如果没有凯文·史密斯和利奥·奥博斯特的努力就不可能开始。他们专业和敬业令人敬佩。还有什么像多个重叠的最终期限能够证明人们的勇气呢，这些家伙就经受了这种考验！

另一个对本书有巨大影响的是我最近三年所从事的工作。我为美国联邦国民抵押贷款协会范尼梅设计了一套 XML 标准用来进行电子抵押贷款，这套标准已经被抵押贷款行业标准维护组织（MISMO）采纳。与范尼梅公司的加里·豪普特曼、詹尼弗·唐纳德和马克·奥利芬特一起工作，那是令人愉快的。与 MISMO 成员一起进行标准修订也是同样令人愉快。与本书关系更为直接的是我在虚拟知识库计划中承担首席架构师的工作。我衷心地感谢 MBI 项目经理丹尼·普罗高和政府项目经理特德·威楚科，他们在整个项目中的支持、敬业的工作以及卓越的管理。特德成功领导智能社区走向知识管理之路。此外，我还要感谢我的架构团队成员凯文·史密斯、乔·拉其库摩和毛利塔·索提斯他们在一系列艰巨问题上的艰苦工作。我也要感谢我在诺斯罗普·格鲁门的团队成员背基·史密斯、马克·利昂和珍妮特·萨金特他们的支持和辛苦工作。最后特别感谢丹尼·普罗高和前任 MBI 高级项目组副主席凯文·艾普斯雷他们的帮助以及对我乔迁亚利桑那州的支持。

还有其他家庭、朋友和相识，他们都对本书有着或大或小的帮助，谢谢你们。我尤其感谢我在麦克唐纳·布拉德利公司的同事，特别是莎伦·麦克唐纳、肯·巴蒂、戴夫·淑萍、盖尔·李斯勒、丹尼·普罗高、苏珊·马来亚、安东尼·萨尔维、乔·布鲁萨德、凯尔·赖斯、戴夫·阿诺德。这些朋友和他们朋友通过他们专业精神、奉献和干劲无论个人或专业上都极大丰富了我的生活。我期待在麦克唐纳·布拉德利公司以后几年更多的挑战和成长。我总是怀着对读者的感激之情。通过最近 10 本书，读者们的感谢、鼓励和更高的要求丰富我的写作经验。我的写作目标从来没变：向读者提供有价值的内容，以可行的、启发的方式讨论困难的课题。我衷心地希望这些目标已经达到，如果没有希望我们的读者能够让我们知道。在此我向读者表达最好的祝愿。

——迈克尔 C. 达康塔

我要感谢我的合著者迈克和利奥。因为他们的辛苦工作，使更多的人理解了语义网的发展。本书是我与迈克合著的第三本，与他一起工作是一件快乐的事情。感谢来自道明尼数字公司的丹·惠勒和来自 CapitalOne 公司的安迪·史特罗斯作为本书的审稿人。再一次

感谢威力出版社的鲍勃·艾里奥特和艾米莉·赫曼，能够与他们一起愉快的工作。我也感谢阿什兰的咖啡，在周六以及周日的下午它们给了我写作的灵感。

虚拟知识库项目工具性地帮助迈克和我关注语义网，能够向政府提供远见性的方案。因为有着特德·威楚科、丹尼·普罗高、克莱·理查森、唐纳德·阿冯德利奥、乔·布鲁萨德、背基·史密斯以及其他人的辛苦工作，我们才可能做这样一个伟大事情。我要感谢格温，她是这个世界最伟大的妻子！

——凯文 T. 史密斯

我向那些鼓励我和支持我完成本书的人们表达我的感激之情，这包括我的同事大卫·法瑞尔、我的妻子克里斯蒂（她完全忍受了我的自我放逐）以及那些匿名审稿人。我应说明本书所表达的观点仅是作者们的个人观点，并不反映 MITRE 企业以及其他公司和个人的官方政策和立场。

——利奥·奥博斯特

# 前言

互联网显著地革新了通过电子方式访问信息的能力。互联网包含大概 30 亿静态文档, 大约有 5 亿国际网民在进行访问。与此同时, 如此庞大的数据使其变得很难发现、访问、展示和维护。这是因为信息内容主要有自然语言呈现出来的。因此, 在致力于解决信息可用性的工具和用方便于人们阅读格式维护的信息之间出现了鸿沟。

为了解决这个问题, 许多新的研究协会和商业企业已经提供机器可处理语义的信息。这些支持对于“提升 Web 网的潜力”是至关重要的。Tim Berners - Lee, 互联网联盟的领导人, 将当前 Web 网的将来看做语义网——一个扩展了的 Web 网, 在那里机器可读的信息和自动化服务远远增强了当前 Web 网的性能。隐藏在数据、程序、页面和其他 Web 资源之中语义的显式表达将激活基于知识的 Web 网, 将提供质量上全新的服务。自动化服务提高帮助人们“理解”更多的 Web 网上内容的能力, 并能对信息源提供更精确的过滤、分类和搜索功能。这个过程将最终导向一个极其知识渊博的系统, 并能提供各种专业逻辑推理服务。这些服务将在我们日常生活的任何方面提供支持, 将使访问信息变得非常普遍、非常必要, 就如同现在我们使用电力设施。

当我的同事和我在 1996 年开始使用本领域“学院派”的协议时, 仅有非常少的网络语言可用。渐渐的后来我们知道还有 W3C 制定的 XML 和 RDF<sup>①</sup> 等网络语言。今天的情形与以前相比已大不相同了。语义网已经作为研究和教育的主题在许多大学建立起来。许多会议、专题学术研讨会和相关期刊也已经开展了这方面的工作。小公司和大公司正意识到该领域对将来公司业绩的潜在影响。然而, 在科学思想向得到广泛应用的技术转换过程中, 仍有一段很长的路要走, 《语义网: XML、Web 服务、知识管理的未来指南》将成为这个转换进程中的一块奠基墙角石。我记起以前曾花了两个月时间来理解 RDF 和 RDFS 究竟是什么。本书将辅助你在两个小时内彻底理解这些技术。本书对于理解语义网、语义网与 Web 服务之间的关系以及应用领域(如知识管理)的潜在优势是非常好的入门指南。本书通过花费你最少的时间, 来帮助你高效的理解这些内容。

因斯布鲁克大学计算机科学学院  
Dieter Fensel 教授

---

① 我记起第一次我问有关 RDF 的问题, 我误听作“RTF”并非常惊讶“RTF”将被作为适当标准用于语义网。

# 目 录

|                           |    |
|---------------------------|----|
| 第一章 语义网是什么 .....          | 1  |
| 语义网是什么 .....              | 1  |
| 为什么需要语义网 .....            | 3  |
| XML 如何用于语义网 .....         | 4  |
| Web 服务如何用于语义网 .....       | 5  |
| Web 服务之后是什么 .....         | 5  |
| 怀疑者如何评价语义网 .....          | 8  |
| 为什么怀疑者观点不正确 .....         | 9  |
| 总结 .....                  | 10 |
| 第二章 语义网的商业案例 .....        | 11 |
| 语义网的好处 .....              | 11 |
| 语义网技术已经存在吗 .....          | 15 |
| 总结 .....                  | 16 |
| 第三章 理解 XML 及其对企业的影响 ..... | 17 |
| 为什么 XML 成功 .....          | 17 |
| XML 是什么 .....             | 20 |
| XML 对于企业信息化的影响 .....      | 31 |
| 元数据为什么不够用 .....           | 33 |
| 语义层次 .....                | 33 |
| 规则与逻辑 .....               | 34 |
| 推理引擎 .....                | 35 |
| 总结 .....                  | 35 |
| 第四章 理解 Web 服务 .....       | 36 |
| Web 服务是什么 .....           | 36 |
| 为什么使用 Web 服务 .....        | 39 |
| Web 服务能解决实际问题吗 .....      | 39 |
| Web 服务会有明天吗 .....         | 40 |
| 怎样使用 Web 服务 .....         | 40 |
| 理解 Web 服务的基础概念 .....      | 41 |
| 如何发现 Web 服务 .....         | 44 |
| 保证 Web 服务的安全 .....        | 49 |
| 安全套接层还不够安全吗 .....         | 51 |
| Web 服务的下一步是什么 .....       | 53 |
| 总结 .....                  | 54 |

|                      |     |
|----------------------|-----|
| 第五章 理解资源描述框架         | 55  |
| RDF 是什么              | 55  |
| 为什么 RDF 没有成为主流       | 63  |
| 什么是 RDF 模式           | 69  |
| 非语境建模是什么             | 74  |
| 总结                   | 78  |
| 第六章 理解其他术语           | 80  |
| XPath                | 80  |
| 样式表族: XSL、XSLT、XSLFO | 81  |
| XQuery               | 85  |
| XLink                | 86  |
| XPointer             | 89  |
| XInclude             | 90  |
| XML Base             | 91  |
| XHTML                | 92  |
| XForms               | 93  |
| SVG                  | 97  |
| 总结                   | 98  |
| 第七章 理解分类体系           | 100 |
| 分类体系综述               | 100 |
| 定义本体谱系               | 106 |
| 主题图                  | 114 |
| 主题图的相关概念             | 116 |
| 主题图对比 RDF            | 121 |
| 总结                   | 123 |
| 第八章 理解本体             | 125 |
| 本体综述                 | 125 |
| 语法、结构、语义与语用          | 131 |
| 本体的逻辑表示              | 140 |
| 知识表示: 语言、形式化和逻辑      | 150 |
| 语义网、基于框架知识表示和描述逻辑    | 150 |
| 当今的本体                | 155 |
| 总结                   | 160 |
| 第九章 形成公司通向语义网的路线     | 162 |
| 典型组织: 淹没在信息中         | 162 |
| 以知识为中心的组织: 我们的发展方向   | 164 |
| 如何实现目标               | 167 |
| 总结                   | 171 |

# 第一章 语义网是什么

“第一步是将数据以机器能理解形式放在网络上，或者将它们转化为这种形式。这样就产生了我所谓的语义网——数据可以由机器直接或间接处理的网络。”

——Tim Berners - Lee, 《编织网络》, Harper San Francisco, 1999

本章旨在解密语义网。本章将解释语义网是现有 Web 网的逻辑扩展，而并非一种遥不可及的可能。语义网既可以实现，也是值得去做的。本书为 Web 网发明者 Tim Berners - Lee 提出的远景提供一条清晰的路径。

## 语义网是什么

Tim Berners - Lee 关于未来网络的远景包括两部分：第一部分是使网络成为一个协作更加深入的媒体，第二部分是使网络可以为机器所理解 and 处理。图 1 - 1 是 Tim Berners - Lee 一份最初的远景图。

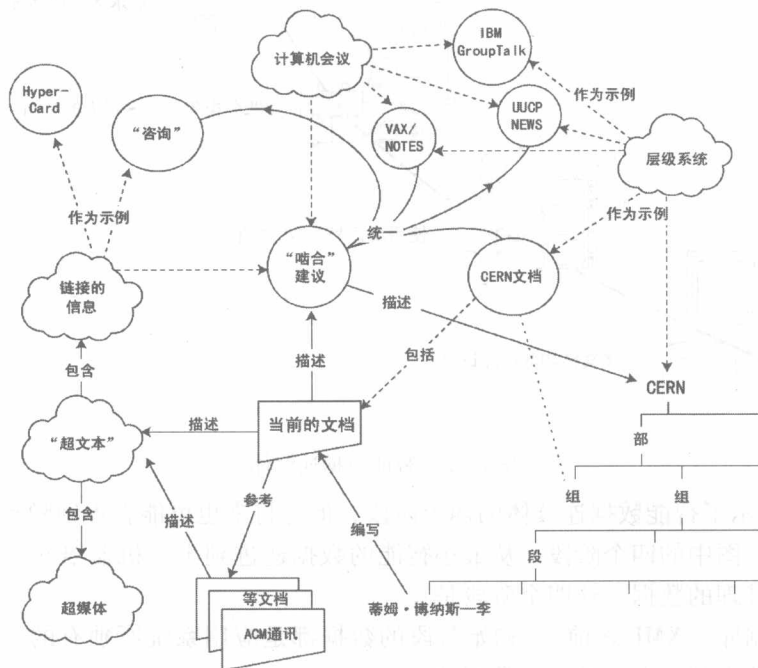


图 1 - 1 CERN 的 Web 网最初版本<sup>①</sup>

Tim Berners - Lee 最初远景所涉及的不仅仅是从 Web 服务器上对超文本置标语言 (HT-

① 著作权 Tim Berners - Lee.