

信息资源

XINXI ZIYUAN

整合技术

ZHENGHE JISHU

彭洁 赵辉 齐娜 编著

010100001011110111000101000100

010100001011110111000101000100

010100001011110111000101000100

0100001011110111000101000100

01011110111000101000100

11000101000100

 科学技术文献出版社

信息资源整合技术

彭 洁 赵 辉 齐 娜 编著

科学技术文献出版社

Scientific and Technical Documents Publishing House

北 京

图书在版编目(CIP)数据

信息资源整合技术/彭洁,赵辉,齐娜编著.-北京:科学技术文献出版社,2008.7
ISBN 978-7-5023-5898-3

I. 信… II. ①彭… ②赵… ③齐… III. 信息管理-研究 IV. G203

中国版本图书馆 CIP 数据核字(2008)第 000941 号

出 版 者 科学技术文献出版社
地 址 北京市复兴路 15 号(中央电视台西侧)/100038
图书编务部电话 (010)51501739
图书发行部电话 (010)51501720,(010)51501722(传真)
邮 购 部 电 话 (010)51501729
网 址 <http://www.stdph.com>
E-mail: stdph@istic.ac.cn
策 划 编 辑 周国臻
责 任 编 辑 周国臻
责 任 校 对 赵文珍
责 任 出 版 王杰馨
发 行 者 科学技术文献出版社发行 全国各地新华书店经销
印 刷 者 利森达印务有限公司
版 (印) 次 2008 年 7 月第 1 版第 1 次印刷
开 本 787×960 16 开
字 数 356 千
印 张 20
印 数 1~2000 册
定 价 45.00 元

© 版权所有 违法必究

购买本社图书,凡字迹不清、缺页、倒页、脱页者,本社发行部负责调换。

前 言

2004年,中共中央办公厅和国务院办公厅联合发布了[2004]34号文件——《关于加强信息资源开发利用工作的若干意见》,该文件首次将信息资源管理问题提到了国家宏观管理的高度,也预示着我国的信息化进入了一个新的阶段。党的十七大报告中,推进信息化的思想贯穿全篇。有关人士认为,随着我国各行业信息化的深入开展,信息化将从物质资源管理开始过渡到信息资源管理。

在信息化建设开始的很长一段时间内,各类信息化建设项目在信息资源方面大多缺乏统筹规划,应用系统之间普遍缺乏标准化的数据接口定义,造成了应用系统林立、系统内的数据/信息不能共享和重复利用的局面,形成了彼此隔离的信息孤岛,信息系统的使用效率大大降低,信息化的快速推进受到了阻碍。如何彻底地消除信息孤岛,有效地整合现有及未来的业务应用系统的信息数据,实现IT与业务的完全融合已成为当前信息化建设的重中之重。

2007年1月,计世资讯(CCW Research)发布的最新研究表明,到2009年,我国74%的大型机构将把信息化建设推进到IT系统整合和业务整合阶段。这个整合阶段又可以细分为5个层次,分别是逻辑整合、物理整合、平台

整合、应用整合和融合。这 5 个层次依次由低至高,每个阶段的整合,都为后一阶段提供支撑,最终实现 IT 和业务的无缝融合。

基于这一认识,我们开始进行了信息资源整合技术的相关研究。本书在广泛调查了各类机构的信息化规划需求的基础上,通过系统的方法梳理和归纳了信息资源整合技术的发展历程、成熟度、技术特性、产品、厂商和市场,建立了一个信息资源整合技术的研究体系。同时通过信息技术成熟度分析等研究,试图帮助用户深入浅出地把握不同技术之间的区别和关联,合理科学地利用好信息资源整合技术,提升信息化水平,最终提升核心竞争力。

全面描述信息资源整合技术,可从整合技术与整合规则两方面来诠释。其中整合技术包括 Seb Wervices、用于信息资源整合的中间件技术、商业智能、协同与搜索技术等;整合规则如 open URL、SOA、SOAP 等等。作为一个较新的研究领域,本书根据当前的热点问题,向读者介绍的是甄选出的几个具有代表性的信息资源整合技术,其他技术、规则或适用于特定领域(例如图书情报领域)技术组的介绍将通过我们研究的不断探索、反思、完善,在今后的修订版中陆续补充。期望未来能呈现给读者一个全面完整的信息资源整合技术的全貌。

本书由 4 部分组成,共分 16 章。第一部分是综述,概括介绍信息资源整合的时代背景,界定整合技术的概念,描述了技术的发展现状与趋势等。第二部分详细介绍本次研究所甄选的 5 种信息资源整合技术,即中间件、协同、商业智能、搜索技术、Web2.0,向读者详细描述了技术的概念与内涵,梳理了技术发展历程,展示了技术的应用领域与典型产品。第三部分是从市场的角度了解技术情况,主要介绍了每种技术相关产品的主要提供商、市场竞争情况及

趋势。第四部分介绍了技术相关产品的选型策略及评估指标体系,分三章(选型策略矩阵、应用关联矩阵和产品选型指标体系)进行阐述。详细描述与其他教材不同的是,本书用了一定的篇幅介绍市场、产品及选型策略与指标,意在让读者更全面、立体地了解在当前时代背景下的信息资源整合技术。

本书主要为进行信息资源研究的专业人员提供研究资料,具有很好的可读性和实用性,可以作为图书馆情报专业的硕士及博士研究生的专业参考教材,也可以用作作为相关机构的培训教材。

本书是集体合作的成果。全书由彭洁、赵辉策划、组织撰稿完成。在本书的写作过程中,我们参考了大量文献,已尽可能列在注释和书后的参考文献中,但其中仍难免有所遗漏。特别是一些从网上下载的资料,由于网页内容变更等原因,无法一一标出来源。这里特向被遗漏的作者表示歉意,并向所有作者表示诚挚的谢意。

在本书编写过程中,孙卫、王惠临等老师为本书进行了审稿,在此谨向他们表示感谢。

编 者

2007年12月

(京)新登字 130 号

内容简介

近年来,我国信息化发展取得了长足进步,信息技术应用与国民经济和社会发展的融合日益增强。随着各种信息系统建设的不断完善,各个系统中累计的各类信息的总量持续激增,信息化建设进入了一个新的阶段——信息资源整合阶段。本书在全面分析信息资源整合技术发展背景的基础上,对信息资源整合技术的概念、发展历程、发展环境和技术成熟度开展深入研究,揭示技术的发展趋势,重点分析了协同、中间件、WEB2.0、商业智能、搜索技术,并对重点技术的厂商和产品进行对比分析,为读者呈现了一幅信息资源整合技术的全景图。本书可作为进行信息资源整合技术研究的重要参考资料之一,亦可作为大中专院校的相关教材。

科学技术文献出版社是国家科学技术部系统唯一一家中央级综合性科技出版机构
我们所有的努力都是为了使您增长知识和才干

目 录

第一部分 信息资源整合技术综述	(1)
第一章 全球信息资源整合的背景	(2)
1.1 信息资源整合的时代背景	(2)
1.2 全球信息技术的发展为信息资源整合提供了可能性	(5)
1.3 信息资源整合现状和问题	(9)
第二章 信息资源整合技术的概念与层次框架	(10)
2.1 信息资源整合技术相关基本概念和内涵	(10)
2.2 信息资源整合技术框架	(16)
第三章 信息资源整合技术发展现状与趋势	(30)
3.1 Hype 曲线原理	(30)
3.2 信息资源整合技术的发展历程	(32)
3.3 信息资源整合技术的发展现状	(36)
3.4 信息资源整合技术的发展趋势	(37)
第二部分 信息资源整合技术的概念与典型产品	(41)
第四章 中间件技术	(42)
4.1 概念和内涵	(42)
4.2 中间件与信息资源整合的关系	(44)
4.3 技术框架	(46)
4.4 技术的发展成熟度及特征	(48)
4.5 应用领域	(52)
4.6 典型产品及发展趋势	(54)
4.7 典型应用案例	(82)

第五章 协同技术	(84)
5.1 概念和内涵	(84)
5.2 与信息资源整合的关系	(86)
5.3 协同技术框架	(88)
5.4 技术的发展成熟度及特征	(94)
5.5 应用领域	(97)
5.6 典型产品	(99)
5.7 典型应用案例	(107)
第六章 商业智能技术	(109)
6.1 概念和内涵	(109)
6.2 与信息资源整合的关系	(113)
6.3 技术框架	(114)
6.4 技术的发展成熟度及特征	(115)
6.5 应用领域	(119)
6.6 典型产品	(120)
6.7 典型应用案例	(131)
第七章 搜索引擎技术	(133)
7.1 概念和内涵	(133)
7.2 与信息资源整合的关系	(134)
7.3 技术框架	(135)
7.4 技术的发展成熟度及特征	(139)
7.5 应用领域	(141)
7.6 典型产品	(146)
7.7 典型应用案例	(151)
第八章 Web2.0 支撑技术	(154)
8.1 概念和内涵	(154)
8.2 与信息资源整合的关系	(163)
8.3 技术的发展成熟度及特征	(165)
8.4 应用领域	(175)
8.5 典型产品	(184)
8.6 典型应用案例	(190)

第三部分 信息资源整合技术的提供商和市场	(193)
第九章 中间件技术	(194)
9.1 主要提供商及其特点	(194)
9.2 市场竞争及趋势	(204)
第十章 协同商务技术	(210)
10.1 主要提供商及其特点	(210)
10.2 市场竞争及趋势	(221)
第十一章 商业智能技术	(227)
11.1 主要提供商及其特点	(227)
11.2 市场竞争及趋势	(239)
第十二章 搜索引擎技术	(246)
12.1 主要提供商及其特点	(246)
12.2 市场竞争及趋势	(256)
第十三章 Web2.0 支撑技术	(261)
13.1 主要提供商及其特点	(261)
13.2 市场竞争及趋势	(269)
第四部分 信息资源整合技术产品选型策略及评估指标体系	(273)
第十四章 选型策略矩阵	(274)
第十五章 应用关联矩阵	(278)
第十六章 产品选型指标体系	(282)
16.1 指标设计依据	(282)
16.2 提出评估指标体系意义	(282)
16.3 指标设计的方法论	(283)
16.4 指标参考的基本理论	(285)
16.5 指标框架	(290)
16.6 细分指标项说明	(292)
16.7 指标计算说明	(294)
16.8 指标应用特别说明	(296)
关键术语	(298)

参考文献	(300)
附录 1 涉及厂商索引表	(302)
附录 2 图表目录	(305)

· 设计加 设计 ·

第一部分



· 设计加 设计 ·

信息资源整合技术综述

第一章

全球信息资源整合的背景

1.1 信息资源整合的时代背景

今天,以信息技术为主导的第三次科技革命正在形成和推进。20 世纪中期以来,人类认识自然的能力以前所未有的速度增长。知识信息的爆炸式增长是信息时代的突出特征,也是信息时代到来的重要前提。信息成为社会经济发展的战略性资源。海量信息的飞速产生,迫使各类机构把工作重心从如何获得信息转到如何准确地过滤和有效利用各种信息上来。信息资源整合由于能有效地解除信息孤岛,提高各种信息资源的利用效率而成为大家广为讨论的话题之一。

1.1.1 信息资源管理成为知识经济时代重要的研究课题之一

1996 年,经济合作与发展组织(OECD)提出“以知识为基础的经济”(The Knowledge Based Economy)的概念,我们普遍称之为知识经济。知识经济是建立在知识和信息的生产、分配和应用上的新型经济,它以知识资源为首要依托,以信息技术为代表的高技术产业为第一产业支柱,其根本特征是知识信息已成为国家生产力发展和经济增长的关键因素。劳动生产率的提高,主要不再靠劳动者的体力和技能,而是靠智力和知识。信息的占

有决定权利的归属,国家间的实力较量也演变成了占领信息科学技术制高点的较量。

目前,世界各国纷纷把发展信息技术作为社会经济发展的一项重大战略目标。西方一些发达国家,通过以信息化为重点的经济结构调整,逐渐由以物质生产及其流通为基础的工业化社会,迈向以信息的流通与积累为主要特征的信息化社会。西方少数发达国家已经进入知识经济社会,我国经过 20 世纪 90 年代初期开始到现在的十几年的加速建设和发展,也逐渐拉开了知识经济时代的序幕。

知识经济的深入发展导致人们的工作性质发生改变——从以物质劳动为主转向以智力劳动为主。信息及其技术如同物质和能力一样作为基本资源,参与社会生产。信息在生产过程中的使用,一方面使产品增值,另一方面随着科技的进步,生产所得的产品也出现了非物质化倾向(现象)。同时,随着信息要素在生产过程中地位的不断提高,也要求劳动力具有更高的知识和技术水平及更高的素质。美国经济学家大卫·伯契运用波拉特的理论,在研究中得出了如下数字:20 世纪 70 年代,在美国增加的 2 000 万个新岗位中,只有 5%属于制造业,而大约 90%属于信息、知识和服务性工作。¹

同时,随着互联网的发展和普及,上网的用户越来越多,也为人们提供了最丰富、便捷的互联网信息资源。目前,连接因特网的国家和地区已超过 160 多个、上网主机 1 亿多台。特别是 web2.0 模式的出现,使延续几十年的人-机交互模式,演变为“人-人”交流模式,个人作者成为互联网信息资源产生的源头,越发加速了互联网信息的大量创生。而且在互联网的环境中,这些创生的信息得以快速地传播,使得人们真正进入了信息大爆炸时代。根据相关统计结果,截至 2006 年 6 月,中国使用和浏览博客的人数达到 6 000 万,拥有博客的人数接近 2 000 万。IDC 最新的研究报告《数字宇宙膨胀:到 2010 年全球信息增长预测》指出,2006 年全球每年制造、复制出的数字信息量共计 1 610 亿 GB,受“富媒体”、用户创建内容和 16 亿网民三大因素推动,到 2010 年,全球数字信息量预计为 9 880 亿 GB。IDC 研究数据发现,如果要将所有制造和复制出的数字信息全部保存,2007 年已经出现了存储容量不够用的情况。

因此,知识经济时代面临着信息资源管理的巨大挑战,信息资源管理的问题已经成为全球性的战略问题之一。

1 王云鹏主编,《信息产业管理精要》,北京:北京邮电大学出版社,2003:3

1.1.2 信息资源管理已经进入信息资源整合阶段

1979年,美国管理信息系统专家诺兰(Richard L. Nolan)通过对200多个公司、部门发展信息系统的实践和经验的总结,提出了著名的信息系统进化的阶段模型,即诺兰模型。诺兰将计算机信息系统的发展道路划分为6个阶段,并强调任何组织在实现以计算机为基础的信息系统时都必须从一个阶段发展到下一个阶段,不能实现跳跃式发展(见图1-1所示)。

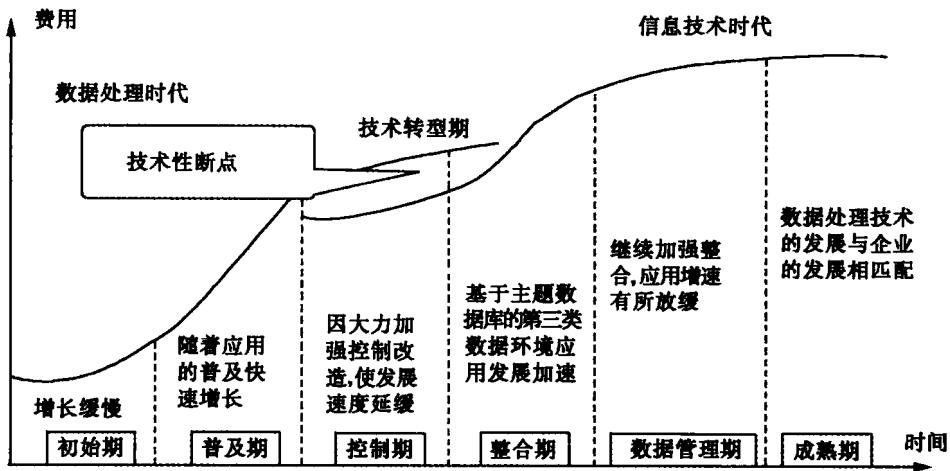


图 1-1 诺兰六阶段模型

第一个阶段是初始阶段。在这个阶段中,计算机刚刚进入企业,一般只在财务部门使用,像管理应收账款和工资这样的数据处理系统。组织中只有个别人具有使用计算机的能力,采购量少,在企业内没有普及。

第二个阶段是传播阶段。信息技术应用开始扩散,企业对软件开发的投入开始大幅度增加,数据处理能力得到迅速发展,但出现数据冗余、数据不一致、难以共享等问题,应用效率不高,有盲目购机、盲目定制开发软件的现象,缺少计划和规划,IT的整体效用无法突显。

第三个阶段是控制阶段。一些职能部门内部出现了网络化,但各软件系统之间还存在“部门壁垒”、“信息孤岛”,信息系统呈现单点分散特点,这时企业管理者已开始从整体上控

制计算机系统的发展,成立了领导小组。这一阶段是计算机管理变为数据管理的关键。

第四个阶段是集成/整合阶段。在控制阶段的基础上,企业开始重新进行规划设计,建立集中式的基础数据库,并建成统一的信息管理系统,IT建设开始由分散和单点发展到成体系。此时增加了大量硬件,预算费用大幅增长,系统更不稳定。

第五个阶段是数据管理阶段。信息成为企业的重要资源。在这一阶段,企业开始选用统一的数据库平台、数据管理体系和信息管理平台,统一数据的管理和使用。各部门各系统基本实现资源整合,信息共享。IT系统规划及资源利用更加高效。

第六个阶段是成熟阶段。此时,信息系统已可满足企业各个层次的需求,从简单的事务处理到支持高效管理的决策,企业真正把IT同管理过程结合,将组织内部和外部资源充分整合和利用,从而提升了企业竞争力和发展潜力。

2007年1月,计世资讯(CCW Research)发布的最新研究表明,到2009年,中国74%的大型机构将把信息化建设推进到IT系统整合和业务整合阶段。一部分中国大型机构的系统整合在2004年就开始了,并逐步向第五、第六个阶段过渡。

1.2 全球信息技术的发展为信息资源整合提供了可能性

全球信息技术为信息资源整合提供了坚实的技术基础。

● 计算机硬件的发展,为信息资源整合提供了基础的物质保证

自1946年第一台电子计算机诞生以来,计算机硬件系统的发展日新月异。20世纪70年代以后,计算机用集成电路的集成度迅速从中小规模发展到大规模、超大规模的水平,微处理器和微型计算机应运而生,各类计算机的性能迅速提高。随着字长4位、8位、16位、32位和64位的微型计算机相继问世和广泛应用,对小型计算机、通用计算机和专用计算机的需求量也相应增长了。单台计算机的计算和存储能力大幅度提高,信息资源首先在计算机硬件水平上不断集成了。

微型计算机在社会上大量应用后,一座办公楼、一所学校、一个仓库常常拥有数十台以至数百台计算机。实现它们互连的局部网随即兴起,整合这些单独的计算机的能力和资源成为了一种需求。

● 数据库系统的发展,为信息资源整合提供了必要的软件支持

第一代数据库的代表是1969年IBM公司研制的层次模型的数据库管理系统IMS和20世纪70年代美国数据库系统语言协商会CODASYL下属的数据库任务组DBTG提议的网状模型。第二代数据库系统是支持关系数据模型的关系数据库系统。20世纪70年代是关系数据库理论研究和原型开发的时代,其中以IBM San Jose研究室开发的System R和Berkeley大学研制的INGRES为典型代表。数据库技术成为实现和优化信息系统的基本技术,也为信息资源的管理和利用提供了可能。

20世纪80年代,随着科学技术的不断进步,各个行业领域对数据库技术提出了更多的要求,第三代数据库应运而生。第三代数据库系统集成数据管理、对象管理和知识管理为一体,具有更好的开放性,支持数据库语言标准;在网络上支持标准网络协议;系统具有良好的可移植性、可连接性、可扩展性和可互操作性等。90年代初期数据仓库的概念被提出并逐步实现产品化,使数据资源进一步整合。

● 网络技术的发展,为信息资源整合利用提供了不可或缺的渠道

1969年,为了能在一旦爆发核战争时保障通信联络,美国国防部高级研究计划署(ARPA)资助建立了世界上第一个分组交换试验网ARPANET,连接美国4所大学。这标志着网络技术的开端,使得计算机与计算机的连通成为了可能。

1986年,美国国家科学基金会NSF资助建成了基于TCP/IP技术的主干网NSF-NET,连接美国的若干超级计算中心、主要大学和研究机构,世界上第一个互联网产生,迅速连接到世界各地。

1990年11月,第一个Web服务器nxoc01.cern.ch开始运行(后改名为info.cern.ch),Tim Berners-Lee在自己编写的图形化Web浏览器“WorldWideWeb”上看到了最早的Web页面。1991年,CERN(European Particle Physics Laboratory)正式发布了Web技术标准。

1995年,IEEE认可了快速以太网,并发布了802.3u 100Bast-T标准,其网络传输速度可以达到100Mbps。基于网络的高速内容传输有了物质保障。

● 服务器的发展,为信息资源整合交互提供了保障

服务器(Server)发展到今天,适应各种不同功能、不同环境的服务器不断地出现,分类标准也多种多样。按应用层次划分为入门级服务器、工作组级服务器、部门级服务器和企业级服务器4类;按服务器的处理器架构(也就是服务器CPU所采用的指令系统)划分