

基于网络环境的信息服务模式及其现实

周 翠 莲

(南华大学 船山学院, 湖南 衡阳 421001)

[摘 要] 基于网络环境下的服务模式是| 种信息集成服务、面向内容的信息服务、 个性化信息服务,正确处理这三种信息服务模式中的知识产权与经济学问题、改善信息系统人机界面、应用网络计量学是实现这些信息服务模式的主要途径。

[关键词] 网络环境; 信息服务

[中图分类号] TP3 [文献标识码] A [文章编号] 1673- 0755(2009)02- 0110- 03

一 基于网络环境下的服务模式

(一) 信息集成服务

信息集成服务是指将大量分散的资源按一定的主题、一定的任务进行智能化的聚集,并通过统一的检索服务,最终实现信息资源 1加 1大于 2使用效果的一种服务。

网络信息资源的最大特点就是信息的差异性、资源的分布性以及管理的自治性。由于不同的网络信息服务系统提供不同的资源与服务,具备不同的功能,人们希望能将这些不同的资源与服务进行集成,形成一个大系统。再说,用户不再满足于通过网络检索出一大堆原始信息,而是要求一个综合集成、媒体丰富以及深度加工的信息,甚至是一个包含知识和解决方案的信息服务^[1]。这就提出了信息综合集成服务的需求。

1.信息集成系统。在信息集成系统中,由于数据管理是由不同的服务器完成的,一旦这些不同的服务器之间需要相互联系,采取一致操作时就有有一个协调问题,如某服务器发生故障,怎么保证不影响整个系统正常运转;又如某个服务器不允许随意删除索引中的记录,那么整个系统的资源更新就需要重新建立一个索引。

在信息集成系统中还有一个查询处理的问题。在分布式的环境下,查询表达式或不同的查询子句将被分配到不同的资源上去,这时同样的一个查询表达式被分配到不同的资源上会产生不同的查询路径,由此会有不同的结果返回,这样会导致查询质量的下降。

网络信息资源量大,种类繁多,管理模式各异,加上某些网络信息提供商的商业动机也将导致信息的误差,因此,需要研究为迷茫、多种类的信息资源提供导航检索功能以及多媒体支持的信息集成服务系统的技术基础构架,具有涉及数

据库的查询路径、数据库的互操作以及数据库的协调管理等问题,从而为分布式的信息资源提供统一的查询界面。

2 资源集成内容。信息资源集成过程中的另一个重要问题是信息内容选择,如数据库的选择,具体涉及如元数据的确定问题、数据库性能与价格问题等。尤其是元数据问题,更是研究的重点,这对于未来的以信息集成为基础的数字图书馆来说尤其重要。

(二) 面向内容的信息服务

尽管网络信息资源庞大,人们查询信息很方便,但查询的质量仍然不尽人意,因此人们将信息服务的研究重点放在了精确信息内容提供上和信息内容的增值处理上。

1. 词表导航

在信息集成环境下,人们要实现跨主题资源的相互操作,实现用户与系统的一种交互式的导航,实现专业指度很高的概念检索,这就用到了词表技术。当今,词表技术研究的一个重要发展趋势就是与网络检索这类应用系统的结合。目前美国伊利诺斯大学与亚利桑那大学正在研究被称为“信息交互空间”(Interspace)的项目,目的是建立一种多媒体信息的语义索引。这种语义索引是建立在一种概念空间(Conceptspace)基础上的,它是基于概念与分类的聚类统计这样一种思路,首先需要从某一网络上的各资源单元中逐步抽取出某一主题领域的相关概念,形成某一主题领域的概念空间词表,然后利用“自组织图像”(Self- Organizing Map, SOM)方法形成概念图,由此形成概念空间。

通过概念空间词表界面,用户可以随意输入任何一单字、词组或多种检索词。这时搜索界面会在一个显示框中显示与用户检索词相关程度排序的概念空间词汇,用户可以从中选择合意的词添加到搜索框中,以提高检索的精确度。

2 自动摘要

[收稿日期] 2009- 01- 12

[项目基金] 南华大学 2008年度教学教改课题《基于教学评估指标的独立学院档案资源建设研究》(编号: 2008GJY 39)

[作者简介] 周翠莲(1964-),女,湖南衡阳人,南华大学船山学院馆员。

从网上自动跟踪索引的大量文件根本无法用传统方式去做摘要,而用户希望在检索结果中为每篇文献提供一个简短的内容摘要以迅速判断文献的相关性,哪怕有点内容的特征和线索也行。日本富士通公司宣称,它已研制成功比现有软件能更准确地从长达 100 页的文件中抽取 1 至 2 页摘要内容,按照题目、主题和其他标准把文件划分为若干章节,并利用关键词判断章节的起始。把文件划分章节后,该软件根据出现的选定关键词判断章节的起始。把文件划分章节后,该软件根据出现的选定关键词分析句子,对每一块分别进行内容摘要。上海交通大学计算中心的“自动编制中文科技文献文摘的实验性系统”也具有相同的性质。自动文摘 (Automatic Abstracting) 从来就是信息内容加工处理自动化的重要课题,已有近 40 年的历史,但是要求自动文摘能够将原文的主题思想或中心内容自动提取出来,具有概况性、客观性、可理解性和可读性,过于野心勃勃,这个难点只有计算机智能化的分析和辨识主题结构的技术得到突破才可望解决^[2]。

3 知识的可视化显示

知识的可视化显示能更好帮助人们对信息内容及其隐含知识进行感知与理解。当前这类研究一个是集中在词表的可视化显示中,另一个则是分析处理的信息可视化显示。

(1) 词表的可视化显示。用于概念空间图像显示的图像词表技术,用于大规模可视化检索与分类检索的自组织图像 (SOM) 技术,并利用这些技术研究开发了各种概念图的实验模型,特点是对概念之间关系的说明采用了三维图形化的表达形式。用户进入该系统界面,通过相应的选择,屏幕会以线形列单形式显示主题概念,选择其中一个,点鼠标右键,系统会显示该概念的注释;点鼠标左键,系统会以三维立体结构形式显示与该词汇相关的不少于 26 个的近邻词。

(2) 分析处理信息的可视化显示。传统的引文统计信息表达往往让人感到那么枯燥难懂。借助于可视化技术,可为传统显示方式注入新的生命力。

(三) 个性化信息服务

所谓个性化信息服务就是针对个人不同的需求提供特殊信息的一种服务。目前,以用户为中心的思想几乎渗透于社会的每个经济活动中。在信息服务中提出个性化服务,乃是人们迫切希望信息资源能够被有效利用的动机所致。

1. 以用户为中心的智能系统

美国国家科学基金会 (NSF) 专门召开了题为“以用户为中心的智能系统 (HCIS) 研究: 信息、交互与智能”的学术研讨会,目标是探讨一个被称为令人兴奋的重要的新领域—以人为中心的智能系统的未来发展。研讨会上人们一致认为这一新的研究可能会产生各种开发工具,促进人们利用计算机,利用今天我们所拥有的信息资源来解决社会与经济问题。与会专家认为首先要认识到我们的资源开发与以人为中心的思想存在着距离,将注意力转向如何为人们提供信息,并改善信息与人之间通信的工具,并且为了使人们能够直接观察与熟练操作信息,还需要对直觉的交互界面进行研究。因为如果没有人机智能交互,超级计算机的巨大功能,超高速的网络,庞大的多媒体馆藏就不能被最有效地使用。

以人为中心的系统必须能够服务于社会每一个人,他们有的可能是计算机使用的行家,有的却是生手。此外,他们年龄还不相同,他们说的语言也不同,甚至在灵活性、视力、听力等方面都有所不同。因此,以人为中心系统的设计者必须要考虑所有可能的用户,对他们每一个人提供服务,不仅能够适应于一般的用户信息消费模式,而且还应该满足个性化的要求。

当前很多的研究已体现了这一思想,比如建立在对用户信息查询需求与查询任务的相关数据研究与总结基础上的搜寻模型,它们可以用来指导用户对整个资源的搜寻,不仅搜寻存在于文献中的信息,而且还要搜寻直接存在于图形、活动行为图表、具有一定含义的图标以及人与系统的请求交互中的各种信息,并还能让用户根据面临的问题与系统进行交互。在这其中,可视化的信息研究,尤其是虚拟环境以及三维图像技术的研究是特别必要的,这有利于人们在解决问题中对信息资源的感知与有效利用^[3]。

2 检索帮助

从个体用户的角度来看,检索帮助导航无疑是满足个人的各种检索需求而提供的一种检索帮助。这时用户需要的是无需花很多努力去学习的检索方法,需要的是易懂易会甚至是十分令人轻松愉快友好的人机交互界面。最近美国的可视化医学检索系统非常成功地开发出了这种界面。

3 推送服务

Push 技术软件开发及应用已在大幅度增长。尤其是在技术与金融领域,对实时信息需求很迫切。美国 Boston 银行就特别重视推送技术的利用,他们利用 P Multi-cast 协议,通过因特网以及后来的内联网广播频道向它的客户提供关键的实时金融信息。通过服务器,这些信息被推送、筛选,最后只将那些个别用户想要看的内容在桌面上进行显示。

利用 Push 技术,企业还可以在企业的内联网上向企业雇员传递公司内部的各种信息。Push 服务提供商还可以为企业用户直接提供技术框架,企业可以利用这个框架建立自己的频道,由此向雇员以及客户传递所选择的信息。比如,营销主管可以接收到有关行业趋势以及竞争者的信息;人力资源部可以将有关工资方面的信息推送到各个办公室的雇员。

Push 技术能提供及时的、个性化的信息。但是,有时也会像一个喷水的水龙头喷出许多用户不需要的信息。关键是如何保持用户信息界面的专业特性,只有那样,用户才能得到有用的数据。

4 智能代理

美国麻省理工学院 (MIT) 媒体实验室软件代理研究小组认为“智能代理是一种能够完成委托任务的计算机系统”,如用于航班预定,在线书店图书订购,从最喜欢的音乐家那里寻找最新的歌曲等。它们不同于普通的软件,可以快速地通过网络寻找我们所需要的各种信息。

代理的智能化机理一致的看法是,作为一种智能代理,它必须具备一种能力,即实时操作与用自然语言进行通信的能力。据此,它们必须能够适应环境,能够适应面向目标的行为,换句话说,智能代理必须要根据用户特定问题进行动

态的操作;智能代理必须要用一种用户理解的语言与他们进行通信,比如告诉他们任务已经完成,或任务被终止了。

人们认为智能代理以及它的自动搜索处理是一种未来的搜索引擎。一个用户可以依赖智能代理技术进行一种搜索,数据库有新记录被添加时,可以命令搜索引擎的代理通知用户。这样一个静态的系统就会变成有效的由用户驱动的动态信息源。有人还描述了未来智能代理在 Cartak 系统中的一个有趣的应用:人们可以在汽车上安装广播中心,里面编有一套用户感兴趣的数据。地方上的经营商可以利用智能代理来接收这些来自驾驶员的广播信息^[4]。比如,驾驶员可以通过广播希望在附近找一家意大利餐馆,或其他什么兴趣。一个合适的提供商代理就会应答,告诉驾驶员在遥远的地方有这样的餐馆,门牌号是多少,有什么样的行驶路线。

(四)数字图书馆

综观全球形势,因特网环境下以拥有学术性信息资源为特色的信息服务业发展的重要趋势就是组织数字图书馆,提供以数字图书馆收藏为基础的智能化的信息集成服务。

数字图书馆含义目前难以统一。但是,数字图书馆肯定不是指馆藏资源电子化后的传统意义上的图书馆,如我国北大图书馆、清华图书馆。数字图书馆应该具有这样一些特点:通过创造一种信息环境来提供获取分散存贮的信息的智能途径,不仅提供电子途径获取原始数据,而且还提供更高级的智能途径获取原始数据中的知识内容信息^[5]。

在美国国家自然科学基金会与欧洲联盟(EU)联合开发国际数字图书馆项目的有关文献中对数字图书馆是这样描述的:数字图书馆代表了一种集计算、通信与大规模的数字内容于一体的新型信息基础结构(Infrastructure)与信息环境。它将有力地支持从电子商务到科技合作领域里的个人与组织对分布于全球的各种知识的利用活动,帮助各行各业接受来自新时代的挑战。数字图书馆的研究目标是将因特网变成一个虚拟的统一的信息空间,为用户提供统一的智能化的知识获取手段。

二 基于网络环境下的信息服务模式实现对策

(一)正确处理信息集成服务中知识产权与经济学问题

知识产权与经济学问题与网络信息集成问题是密切相关的。比如网络资源开发中的原始数据的利用、复制与获取问题对于数据的索引制作、元数据及其登记是很重要的,而

在互操作过程中,这些问题都将涉及到知识产权与经济学问题。目前人们更加关注的是对以信息集成服务思想为基础的全球性数字图书馆领域里的知识产权与经济学问题。

(二)改善信息系统人机界面

对于网络信息集成系统来说,需要研究如何提供一个便于用户使用的统一检索界面的问题。如目前人们正在研究举例提问、自动提问扩展以及交叉参考提问等技术,由此来实现用户提问帮助,构建一个有针对性的搜索结构。检索结果的可视化表达也是当前研究的一个重要内容。通常检索结果都是按文献清单排序显示,现在人们试图用图像形式展示枯燥检索结果,从而使被检索的文献内容能被更好理解。

(三)应用网络计量学

自从1997年人们首次提出“网络计量学”(Webometrics)以来,人们已将其应用于信息资源内容开发和服务的研究领域。人们应用网络计量学对因特网站点的数量、超文本节点链接的深度、科学文献的标识和特征、文件的延伸、引文的数量和结构、关键词的命中率、用户点网页的频率、各种类型信息的地理分布、各种语言出现的频率等方面进行统计分析,建立数学模型、网络图表和矩阵图,以达到找出Web信息源,分析用户的检索行为和引文形式或行为,评估网络信息的即时响应和网上科技期刊的影响因素,分辨科学的知识结构,考察搜索引擎的检索功能和稳定性,开发功能更强的搜索引擎和服务软件等目的^[6]。

[参考文献]

- [1] 郭佳慧. 信息集成服务研究概述[J]. 图书馆学刊, 2009(2).
- [2] 栗久珍. 网络论坛情报收集系统研究[J]. 现代情报, 2007(12).
- [3] 李 玮. 论数字档案馆的个性化信息服务[J]. 法制与社会, 2009(2).
- [4] 梁启娟, 秦 斌. 浅谈智能代理在数字图书馆中的应用[J]. 图书与档案, 2009(1).
- [5] 朱凤荣. 数字图书馆是网络时代图书馆发展的必然趋势[J]. 黑龙江史志, 2007(1).
- [6] 邱均平, 张 洋. 网络信息计量学综述[J]. 高校图书馆工作, 2005(1).

On Info Service Models in Context of Internet and Their Implementation

ZHOU Cui-lian

(University of South China, Hengyang 421001, China)

Abstract Service models in the context of internet refer to an info-integrated service, a content-targeted service and an individualized info service. These models are implemented mainly by a proper treatment of the intellectual property and economy in these services, a betterment of the human-computer interface in the information system, and an appropriate application of the network technology.

Key words Net Context Info Services