

刘明晶

数字化网络时代的图书馆管理系统探索

摘要 在对现有图书馆自动化建设存在的问题进行了分析的基础上,提出未来图书馆的系统模式,即将文献资源作为一种知识/信息存储和传递工具资源,利用现代网络和信息处理技术将文献资源进行有效配置,以实现价值最大化。图1。参考文献11。

关键词 数字图书馆 知识抽取/管理 个性化服务 数字化网络

分类号 G250.73

ABSTRACT Based on an analysis of present situations of library automation, the author proposes system patterns of future libraries, which means regarding document resources as resources of knowledge/information storage and transmission and effectively allocating document resources by modern network and information processing technologies to maximize values. 1 fig. 11 refs.

KEY WORDS Digital library. Knowledge extraction/management. Personalized services. Digital network.

CLASS NUMBER G250.73

1988年初~1990年底,我曾作为软件负责人参与了文化部“集成图书馆自动化系统(ILAS)”最初的研制开发工作,在ILAS系统开发中采用了基于LDBMS(图书馆专用数据库管理系统)为基础的软件体系结构。经过了10多年的改进和发展,该系统已进入成熟期,成为同类系统中的佼佼者。

在新的世纪,图书馆自动化系统如何发展,成为图书馆自动化研究人员关心的问题。

文章将阐述关于未来图书馆管理系统的核心功能和软件体系结构,以及系统应用说明。是笔者多年来观察、分析和思考的结果,希望能对新一代图书馆自动化管理系统的研制有所裨益。

1 核心功能表述

本文关于未来理想图书馆系统功能的定义出发点是依据人们使用图书馆的最终需求,试图直入主体、减少中间的低效率过程。

目前我们所见的文献资源只是数据、信息和静态知识的一种简单的载体。新系统的核心功能要求是充分结合读者资源与社会资源,对文献资源进行合理的一体化考虑配置方法。对此,本文归纳了4点,如图1所示。

下面对图示的4个核心功能分别进行解析。

1.1 获取

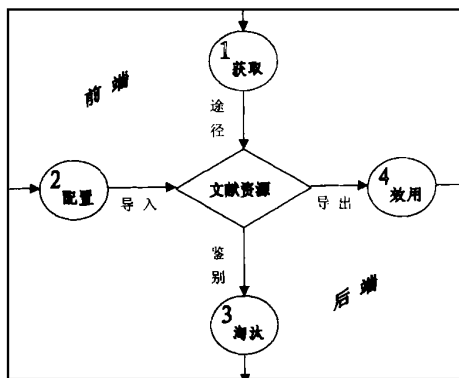


图1 管理系统核心功能

图示说明:

针对文献资源,前端是获取和配置过程,后端是淘汰和效用过程,从前端到后端,这是一个直接对应过程,从后端到前端,这是一个指导过程;

4个核心功能的作用都与读者密切相关,这是把未来图书馆的建设和发展权利部分地转到了读者手中;

4个核心功能是相互联系,密不可分的循环运动的价值链,每个功能所产生的价值都来自于整体的关联和运动结果。

文献资源的获取是指针对不同的对象,采取不同的措施和途径,以获取更广泛、真实、全面的与图

书文献相关的原始数据和信息。获取功能主要强调对数据和信息资源提供者的分类及管理。

一本图书,它的直接资源提供者是该书的作者。目前,数字图书馆是将作者的简介及作者对图书所做的一般性推介在网上进行展示,这远远不能满足读者的需求。新系统要做的就是在网上为读者与作者建立沟通的桥梁。读者根据需求通过系统获取信息或在网上与作者直接对话;系统对读者建立信息库,以便对获取的信息进行管理。

除了作者外,文献资源的提供者还包括领域专家、跨领域专家、直接技术使用者等。

领域专家:对于涉及到研究领域的文献所传播的本专业的知识结构进行批判性地改进或在网上表述自己的看法和意见,系统同样进行全面记录和管理。

跨领域专家:一个领域的专业研究人员,同时具备其它领域的专业知识和通用基础能力,而对其它领域的知识不具备全面批判的能力,但他们具有提出改进和发展的潜力,因此,系统也重视他们的看法和意见。

直接技术使用者:是实践和使用知识的人员,他们通常具有提出和验证的能力。对这一广泛群体,系统将提出众多的吸引方法,激励他们提供信息。

1.2 配置

在获取文献资源的基础上,新系统将对资源之间的组合关系进行新型的智能化信息处理,例如:图书与图书之间、作者与作者之间,按照图书内容的关联性,组合不同作者与图书的资源,为读者提供全新的、符合其需要的针对性信息,减少搜寻时间,获得充分信息。

以下是三种有效的资源组合方法:

关联存储:新系统能够通过学习每一类资源之间的关联来进行工作。通过行为选择判断,当两位读者同时对某一本书发生兴趣时,系统会在两者之间建立联系,进行归类;当同一类别的读者达到一定数量后,新系统就可以建立标本模式,在关联中发掘他们的行为规律并进行存储,进而指导新的读者。

联想记忆:新系统能够按内容存取记忆。当网络输入任一信息时,系统都能够给出一个联想的响应状态,当网络输入不同的标准模式时,网络能够按照某个标准通过联想来输出或输入最为相似的标本模式。

推理意识:新系统自动发展了寻求潜力规则所需要的资源组合方式,通过积累和验证,逐步建立类似于人的推理意识,主动对可能关联和联想的内容进行行为组合。

以上三种资源组合方法是递进的,从关联存储到联想记忆再到推理意识,遵循着由易到难的发展顺序,隶属于不同的功能实现阶段。

1.3 淘汰

淘汰功能的实现依赖于新系统所提供的综合判断。

综合判断是指系统在获取、配置、效用等功能实现过程中,通过建立起来的评判标准,对一些已经证明和暂时证明无效的数据、信息,进行专门存储,同时对其价值属性进行分类记忆。

综合判断包括记录各个功能应用结果,得出有价值信息和知识,并对之进行分类存储和管理。新系统寻求可应用于服务中的概念和规则,进而简化整个淘汰过程。第一,给相似的、反复性的情况提供思想以产生可预见的影响;第二,给新形势提供构思以开展革新;第三,从反复的应用和实证中提炼指导原则。

1.4 效用

效用是指读者能够从新系统所提供的服务中获得好处(价值导出),并在此基础上促进新系统发展。效用功能的实现主要包括以下三个阶段:

第一阶段:读者资源和需求分析。

第二阶段:将新系统配置资源以读者能够明白使用的方式交给读者;

第三阶段:提供指导,建立有效的交流反馈机制。

新系统在提供基本指导的基础上,重点为读者提供导读、知识经验指导和系统积累功效的传播。系统对记录不同读者的不同应用结果进行差异比较,发掘因读者自身原因而产生的不同理解和使用方法及系统原因存在的指导表达问题,增加系统积累内容,改进系统功能,使之更加合理。

效用产生来源于新系统对人、内容、技术的综合,读者能够取得学习、识别、记忆等效用的直接推动力是系统所提供的指导。

2 管理与整合

未来新系统所提供的个性化指导要求更有针对性及实效性,这是软件功能设计的重心。建设的起

点要求是通过对人脑的形象思维、联想记忆等进行模拟和抽象,实现与人脑相似的学习、识别、记忆等信息处理能力。新系统特性要求与人结合,即对机器模仿较差的问题通过人的作用使机器做出改进。另外,建立知识库,除了能够获取增加的信息处理能力外,还能使系统不断地改进工作。

2.1 基本管理

基本管理一方面对文献和读者分别进行管理,另一方面对文献资源关系群体之间建立的关系进行管理。基本管理以读者为中心,通过新系统的统计分析,提高文献的利用率,对读者提供基本的指导服务。

基本管理主要包括 网上数据挖掘系统:收集访问数据,对文献、读者信息进行统计分析,掌握用户服务需求,提供符合用户期望的服务风格和内容。

建设流程的控制和支撑系统:确保内容生产的速度和质量。在建设过程中应用科学化管理,对文献、读者以及它们之间关系进行板块分割,提高各板块的利用率。如:当读者第一次进入数字化图书馆,系统在分析判断过程中会以询问方式建议他输入需求信息或要求他同时提供相关的个人信息。在整个过程中,读者可以随时中止询问,而系统亦会根据已掌握的信息提供个性化的导读服务,按照时间要求排序,使读者作出的选择简单化,一步到位。并根据读者的兴趣,提供文献说明,了解系统导读工作的过程,以获得读者的信任。

个性化服务与读者信息数量有关,系统记载读者信息越多,服务就越具有针对性。

当读者第二次进入数字化图书馆时,系统会自动对读者进行个人鉴别,并主动提供与读者需求相关的信息。这些信息需要保持一定的变动性,才能对读者产生持续的吸引力。

系统要做的就是将这些关系关联起来,记录这些关系的变动情况,并及时提供给读者。而读者将从兴趣入手方便、高效地进行学习并掌握知识,这样所提供的信息就得到了最大化的利用。

2.2 知识管理

知识管理是在对各种来源的信息进行搜集、整理和分析的基础上,对智力资源进行最大限度的利用。针对不同的读者提炼文献中有价值的信息,通过建立知识库,延伸和扩大服务。

知识管理的应用软件主要有:客户关系管理(CRM):识别、确定在市场竞争中最有利可图的目标

市场和目标客户群。通过读者的需求、阅览习惯、首选的交互销售和生活方式的选择进行分析,创造一种更有效、更个人化的商业关系。专家系统:通过人工智能分析和管理的,扩展市场商业价值及未来图书馆的内部知识管理。主要有:知识库、推理引擎、解释工具、知识获取工具、用户界面等。

构成知识库的知识有两类:第一类是领域的真实,是广泛共有的知识,也就是写在书上的知识和常识;第二类是探视式的知识,它是一个领域中正确的实践和推理所需要的判断性知识。如:假设未来的图书馆里的书籍分为三种:一种是图书馆自有的;第二种是作家自己提供的作品;第三种是读者提供的自己喜爱的其他人的作品。三种书籍通过系统建立了不同的知识单元,针对每一种书籍,系统提供的导读推荐也不同:第一种书籍的导读是由系统综合多个作家(专家)和读者信息作出,它需要逐步积累;第二种书籍的导读一般是由作家一个人作出,可以一步到位;第三种书籍的导读通过系统可以综合多个读者信息作出,它的价值是随着信息数量的提升而提升的。

读者可以通过图书馆将自己所拥有的知识贡献出来,系统将对不同的读者、不同的知识进行归类 and 记载,对于不同的知识利用给予相关的价值回报。

读者在数字化图书馆中不能借阅的书籍,通过系统展示,也可以促进书的销售,为作家带来经济效益。

对于大多数可以借的书,与传统不同的是,读者可以不通过图书馆而直接在读者之间相互借阅。读者只需在交换前在系统中进行登记,由系统跟踪书去向和确定还书期限,到时读者就会借到所需要的书。

2.3 从分散到整合

分散是指将基本管理和知识管理中所用的系统软件、配套的数据库管理系统及相应软件,进行分散决策,产生不同作用的渠道,对新的机会及服务需求作出迅速反应。

整合是指将基本管理和知识管理综合利用,在应用中形成一个如核心功能中所描述的价值链。

从分散到整合,主要采用网络技术、信息技术、数据处理技术及相应设备作为软件开发工具,这是核心功能运用所必需的技术支撑。在读者看来,新系统是一个整体,它不会因面对不同子系统的要求而茫然。读者和图书馆自身需要的满足,来自于不

同子系统对其需求的整合,它不是单一系统作用的结果,而是整体运用的结果。

3 应用说明

3.1 应用模式

新系统的应用将采用统一的模式,因为读者及文献资源关系群体存在地区上的差异,收集到的数据和信息也不同,所以新系统通过知识管理以及分散与整合产生知识经验,从而满足各种需求,并带动自身的发展。

3.2 新系统开发

本文提出的新系统是一个综合性软件,通过它的核心功能开发及软件体系建立,实现文献资源的科学管理和图书馆的应用功能扩展,对用户需求产生及时反应,并提供逐步深入的个性化服务以切合发展的需要,保证新系统的开放性和可持续发展。新系统技术指标的要求是先进性、开放性、积累性、发展性;管理指标要求是达到效率、增值的评判标准,充分促进交流和沟通。

3.3 应用目标

充分利用虚拟和实际两种文献达到资源共享;了解未来社区的多元需要,全力推动和支持科技化,使图书馆真正做到“无墙壁”限制;通过对读者、文献等信息资源的掌握,产生有价值的信息知识,使应用更加具有预见性和更为有效;使未来图书馆的建设者和参与者均取得收益,除了以读者为主要服务对象外,还要极大地推动企事业单位的发展,协助它们适应变化需要积极变革取得增值;充分利用软件价值,提供全方位的个性化指导,提高整个社会的认识和实践水平,带动整体经济的发展。

参考文献

- 1 Budi Yuwono, Dik Lun Lee, *WISE: A World wide web Resource Database System*, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 1041-4347/96 \$05.00 ? 1996 IEEE Vol. 8, No. 4: AUGUST

- 1996: 548 ~ 554
- 2 Jur-Tae Kim, Dan I. Moldovan, *Acquisition of Linguistic Patterns for Knowledge-Based Information Extraction*, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 1041-4347/95 \$04.00 ? 1995 IEEE Vol. 7, No. 5: October 1995: 713 ~ 724
- 3 Paolo Liberatore, Marco Schaerf, *Arbitration (or How to Merge Knowledge Bases)*, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 1041-4347/98/ \$10.00 c 1998 IEEE Vol. 10, No. 1: JANUARY/ FEBRUARY 1998: 76 ~ 90
- 4 Dieter Fensel, Jürgen Angele, Rudi Studer, *The Knowledge Acquisition and Representation Language KARL*, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 1041 ~ 4347/98/ \$10.00 c 1998 IEEE Vol. 10, No. 4: JULY/ AUGUST 1998: 527 ~ 550
- 5 Hiroko Fujihara, Dick B. Simmons, Newton C. Ellis, Robert E. Shannon, *Knowledge Conceptualization Tool*, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 1041 ~ 4347/97/ \$10.00 ? 1997 IEEE Vol. 9, No. 2: MARCH-APRIL 1997: 209 ~ 220
- 6 保罗·萨缪尔森等. 经济学. 北京: 华夏出版社, 1999
- 7 Ralph M. Stair 等. 信息管理系统. 北京: 机械工业出版社, 2000
- 8 Abhijit S. Pandya 等. 神经网络模式识别及其实现. 北京: 电子工业出版社, 1999
- 9 [美] Sybase 公司. Sybase 白皮书: 可分析的顾客关系管理. 1999
- 10 (美) IBM/ Notes 公司. Notes 技术白皮书: Lotus 与知识管理. 1998
- 11 刘明晶. 人类新商业——web 经济群落. 深圳: 海天出版社, 1998

刘明晶 北京中体网科技发展有限公司技术总监, 文化部文化科技开发中心自动化研究所兼职研究员。曾参加文化部“集成图书馆自动化系统(ILAS)”的研制并担任软件组组长。通讯地址: 北京市天坛东路 88 号网联大厦 604。邮编 100061。

(来稿时间: 2000-07-28)