

# “互联网 + ” 时代医学高校数字图书馆知识发现系统研究

刘红丽

崔 静

(山西医科大学图书馆 太原 030001)

(山西省科学技术情报研究所 太原 030001)

〔摘要〕 介绍医学高校数字图书馆知识发现系统的体系结构, 包括知识服务交互层、知识发现处理层、数据挖掘传递层、基础数据资源层, 探讨知识发现支撑下医学高校数字图书馆的知识服务模式, 即基于嵌入式、基于文献数据库、基于协同方式的知识服务模式。

〔关键词〕 “互联网 + ”; 医学高校; 数字图书馆; 知识发现

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2017.05.003

**Research on Knowledge Discovery System of Digital Library of Medical University in the "Internet +" Era** LIU Hong-li, Library of Shanxi Medical University, Taiyuan 030001, China; CUI Jing, Institute of Scientific and Technical Information of Shanxi, Taiyuan 030001, China

〔Abstract〕 The paper introduces the system structure of the knowledge discovery system of digital library of medical university, including knowledge service interaction layer, knowledge discovery processing layer, data mining & transmission layer and basic data resource layer, and discusses the knowledge service mode of the digital library of medical university supported by knowledge discovery, namely, the knowledge service mode based on embedded system, document database and collaborative mode.

〔Keywords〕 "Internet + "; Medical university; Digital library; Knowledge discovery

## 1 引言

“互联网 + ” 即 “互联网 + 传统行业”, 具体来说就是以互联网作为交互平台, 大量采用最新的技术手段 (如云计算、物联网、移动互联网、大数据等) 来促进互联网与传统行业的无缝链接, 以实现行业新业态。2015 年 “互联网 + ” 上升为国家战略, 自此以互联网为标志的信息技术革新彻底改变

了社会政治经济发展创新模式和人类生活方式。在医学教育领域, 随着 “互联网 + ” 概念的提出, 医学教育改革从学制改革、课程改革到案例教学、翻转网络课堂、基于问题的教学法、慕课等广泛融合与应用, 使高等教育逐渐由封闭式院校教育走向更加开放广阔的领域, 而这为学术研究、医学教育和临床教学提供信息服务的医学高校图书馆提出了新的更高要求。医学高校图书馆在信息技术化以及医学发展超前化的知识环境下, 迫切要求数字化、智慧化, 开展以用户为主的服务模式, 解决特定的需求和问题, 实现动态的个性化、主动性服务。

欧洲研究图书馆协会发布《数字时代知识发现海牙宣言》, 指出内容挖掘、数据抽取工具是数字

〔修回日期〕 2017-04-17

〔作者简介〕 刘红丽, 硕士, 助理馆员; 崔静, 硕士, 实习研究员。

时代知识发现的关键<sup>[1]</sup>。在法国里昂召开的第 80 届国际图书馆协会联合会 (International Federation of Library Associations and Institutions, ILFA) 大会提出了图书馆要用知识汇集的思维来发展服务<sup>[2]</sup>。医学高校数字图书馆作为可扩展的知识网络系统中心,在揭示内容与提供服务的过程中,提倡以用户为中心来实现知识最大化增值和智能化重组。而知识发现则用于发掘用户个性化特征以及知识资源关联,倡导用户参与,对用户的浏览轨迹进行记录和关联分析,辅助优化馆藏结构、学科建设以及用户的知识性服务。由此可见,构建医学高校数字图书馆知识发现系统可以为读者提供更深层次的知识服务,更好地实现师生个性化的生物医学科学研究。

## 2 医学高校数字图书馆知识发现系统构建

### 2.1 概述

医学高校数字图书馆知识发现系统是一种基于 Web 的面向全校师生的服务,将分散在不同地理位置的同构、异构医学生物科学数据利用分布式计算技术进行集中分析处理,发现知识的信息资源服务集成系统。该系统是由馆员、师生用户和信息服务代理构成,通过需求发布、检索查找和推送知识等交互操作,达到各种资源之间的协同工作,为用户便捷快速找到相关符合需求的文献,最后实现知识共享。国内外数字资源供应商利用互联网技术研发的基于知识发现的知识服务系统平台在图书情报领域已经形成,如 Summon 知识发现系统的“一站式”检索平台、超星学术资源发现系统等。这些知识发现系统内容包括图书馆本身的数字馆藏、自建数据库以及订购的各类远程数据库电子资源,可为用户提供统一资源发现与获取的多种服务<sup>[3]</sup>;但是系统没有设计有关研究用户知识需求行为的功能,忽视了利用用户知识行为来对数字图书馆的资源进行深度挖掘。按照笔者对医学高校数字图书馆知识发现系统的前期研究,完备的体系结构应该具备对生物医学数字资源和对读者知识信息行为进行深度挖掘分析的能力,还应该建立用户信息和资源之间的联系,对用户的知识行为方式进行规划和描述,

为其定制个性化的推送服务。该系统主要利用分层结构构建核心模型体系,分为 4 个层次,从上到下依次为知识服务交互层、知识发现处理层、数据挖掘传递层和基础数据资源层,见图 1。

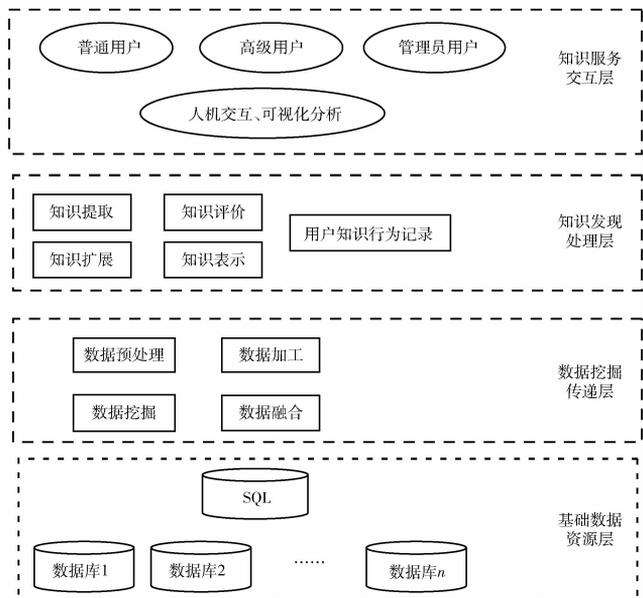


图 1 医学高校数字图书馆知识发现系统体系模型

### 2.2 知识服务交互层

知识发现系统的最顶层,包括用户和可视化交互界面,主要实现知识图谱可视化分析和人机交互功能。系统用户根据权限不同分为普通用户、高级用户和管理员用户。普通用户通过移动终端设备下载图书馆 APP,使用系统提供的图书借阅查询和咨询服务;高级用户是图书馆咨询馆员,通过知识服务交互层对信息进行分析处理,为普通用户提供个性化的优质服务;管理员用户通过知识服务交互层对系统的运行进行维护和更新升级。知识服务交互层是知识发现系统与用户接触最直接的平台,根据用户检索操作规律形象展现信息数据资源的分析结果,给用户可提供分析的拓展信息数据并更深层次挖掘隐藏的知识规律,协助高级用户更好地处理信息。如果高级用户反馈的内容不符合普通用户的需求,该层会以可视化的形式提示用户再次发起请求,再对更深层次的知识进行挖掘,反馈给高级用户使用。

## 2.3 知识发现处理层

知识发现系统的核心层，其主要功能是系统的组织经过分析和深度挖掘数据。该层对知识发现要解决的问题进行了详细定义，制定了一系列编码体系，利用内容分析方法和挖掘算法对数据进行知识分析编码，即对信息进行预处理，将处理好的信息用向量表示，根据相似算法和聚类算法实现信息串并处理和自动聚类，将信息转化为知识并对其进行提取扩展、评价和表示，将发现的有价值的知识信息存放至知识库。知识发现处理层分析挖掘检索系统会自动记录用户的有关检索轨迹，得到用户的职业、性格或兴趣需求以及用户对某类知识的偏好等，从而构建出用户知识行为数据库，包括用户个人信息库和知识行为记录库。知识库和用户知识行为数据库是整个知识发现系统的重点和关键。高级用户应用综合分析方法通过知识发现处理层提供知识分析服务，利用文献资源之间的相关性多角度分析挖掘有效知识信息，为用户科学研究服务。此外，知识发现处理层还可以通过对普通用户的知识行为进行分析处理得出用户行为偏好，生成报告反馈给高级用户。

## 2.4 数据挖掘传递层

知识发现系统最基础的组成单元，主要功能包括数据预处理、数据加工、数据挖掘、数据分类和数据融合等。数据预处理是对要分析的数据进行消除噪声、重复清理、集成变换等操作。数据加工对处理过的数据进行再加工，找到有用信息。数据挖掘过程是依据用户问题与加工数据匹配算法，找出隐藏于数据中具有关联性、结构性、差异性和规律性的数据。数据分类根据结构化和非结构化等方式进行分类，或是根据数据种类和用途不同等方式进行分类，同时也考虑数据的异构性<sup>[4]</sup>。数据融合对多源数据进行集中整理，利用数据处理融合技术将不同站点、不同种类的数据进行整合处理，展现数据的整体特征。通过对数据的挖掘分析，揭示数据的内涵以及语意的认识层次，得到处理效率更高的数据，采用数据传输协议进行传输，更好地为知识

发现处理层所应用。

## 2.5 基础数据资源层

位于整个体系结构的最底层，是知识发现系统运作的基础。该层由许多数据库组成，包括医学高校数字图书馆本身所有的各类数字资源，如情报资料、用户个人信息、用户需求信息、知识库信息等，相当于一个大的数据库。这些数据集合主要来源于用户利用 SQL 语言查询进行检索过程中被捕获的关联性数据、用户注册、在线交流以及用户知识行为特征等数据。

# 3 医学高校数字图书馆知识发现系统服务模式

## 3.1 概述

在医疗信息化建设过程中，医学高校图书馆建设也是很重要的一部分。医学高校数字图书馆的信息化过程是利用互联网技术构建知识发现系统，该系统的服务模式以用户为中心，挖掘出满足用户需求的有用知识，使数字资源利用率最大化，体现了知识找人的目的。具体来讲其服务模式的主要特点是：用户在对生物医学知识集合进行检索时，知识发现系统自动将与之相关联或者相似的知识推送给用户，让读者充分利用图书馆内与之需求相符合的知识，见图 2。

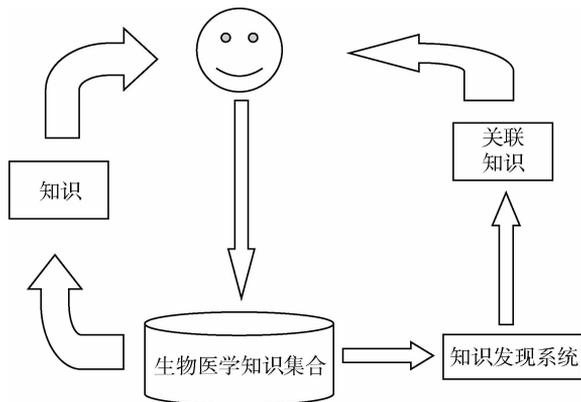


图2 医学高校数字图书馆知识发现系统的服务模式

### 3.2 基于嵌入式知识服务模式

医疗信息化建设构建了一个大型的数字医疗平台, 医学高校图书馆利用知识发现系统将服务整合和嵌入到数字医疗平台中, 传统服务将升级为嵌入式服务模式。医学高校图书馆嵌入式服务主要为临床医学研究以及护理和基础医学研究提供学科服务, 专业馆员将馆藏信息嵌入 QQ 群、微信群等社交网络上, 快速把握用户的信息需求。用户可通过各种移动设备利用图书馆资源和服务。如华盛顿大学图书馆提供的嵌入物理空间的信息服务, 指派生物信息专家进驻医学院和公共健康学院实验室中工作, 为研究人员提供核酸序列分析等软件支持服务, 为院系科研人员提供信息素养教育, 解答咨询问题<sup>[5]</sup>。Welch 医学图书馆嵌入式信息服务, Welch 空间延伸图书馆的物理空间到各院系之中, 将服务嵌入到科研和医疗的工作流中, 保证用户可以随时随地获取信息<sup>[6]</sup>。嵌入式知识咨询是图书馆知识发现系统对知识进行增值和智能重组后, 利用该系统以及简易信息聚合 (Really Simple Syndication, RSS) 推送, 更好地定位用户的个性化服务, 根据挖掘分析的用户行为需求特征, 为用户提供满足其知识需求的相关知识信息。图书馆员应在基于嵌入式知识咨询服务中重新定位其职能价值, 以学科联络专员、信息咨询专家、信息资源建设者以及专业信息素养教育者的身份进行服务, 主动了解用户需求, 根据用户的需求搜集整理资料, 做到以用户为中心, 围绕医学科学发展的需求进行个性化知识组织和重建, 使嵌入式服务进一步促进医学科学的发展。

### 3.3 基于文献数据库知识服务模式

基于文献数据库提供知识服务, 其目的是启发科研思路, 辅助科研决策。专业馆员以 Web of Science、PubMed 及中英文医学专业文献数据库等为知识库, 针对文献量的时间改变情况, 整理出用户检索文献领域的整体发展趋势, 深入挖掘分析高引用和高影响因素的文献, 找出知识融合与分散的演化

路径。在进行文献分析服务时, 可以利用集成有共词分析、抽取或聚类分析医学实体关系的生物医学知识发现工具, 如希腊生物技术研究中心提供的医学主题词表主题词关联关系分析挖掘工具<sup>[7]</sup>、德国计算生物学与数据挖掘研究中心研发的文本挖掘工具 Alkemio<sup>[8]</sup>。或者自建数据挖掘平台, 针对学科开展知识发现服务。该模式可为医学高校科研人员提供大量已发表文献中隐含的科学知识, 简化其阅读文献的过程, 辅助用户进行医疗诊断分析, 指明科研决策的发展方向。

### 3.4 基于协同方式知识服务模式

基于协同方式提供知识服务是指图书馆学科馆员与学科信息专家利用知识发现系统实现知识共享的协同信息行为, 以及与医学院系科研管理员经过协同而得到知识, 将这些数据存放在协同知识库中, 再对数据进行序化整理后转移到共享库, 以此类知识来满足用户的科学研究需求, 增加图书馆知识储备, 使协同知识得到共享。基于协同方式知识服务模式是医学高校数字图书馆数字资源经过知识外化和知识内化过程, 以实现知识共享机制。知识外化过程是学科馆员与医学院系科研管理员之间利用知识发现系统进行协同合作, 学科馆员和学科信息专家根据自己的检索技能利用知识发现系统自动定期推送与科研项目研究领域相关的学术动态以及同行研究进展, 掌握一定的科技评价工具与方法, 构建科学合理的综合指标体系, 为科研项目成果提供绩效评价服务; 同时, 医学院系的教学科研管理员将图书馆学科服务目的、形式、内容等对院校所有学员做大力宣讲及培训。而知识内化过程是学科馆员与学科信息专家之间建立知识协同机制, 让知识在相互学习与借鉴中得到分享与转移。知识协同模式能促使难以规范化表达的隐性知识显性化, 对产生的零散庞杂的显性知识进行分类和整理, 使之成为新的显性知识, 加以利用来实现个体知识的团队共享, 这一知识转化过程为知识共享创造了条件, 达到知识的创新与增值, 提高学科馆员的工作能力和水平。

## 4 结语

医学高校数字图书馆作为学校的文献情报中心,在医学高校教学和科研方面扮演着非常重要的角色。医学高校数字图书馆紧跟生物-心理-社会的医学发展模式,构建的知识服务系统有效提高了数字资源的利用率,帮助用户拓展知识、拓宽思维,主动地根据用户的科研教学需求,增强了数字图书馆对潜在知识的挖掘能力,提供即时、有效、有深度的信息服务与知识服务转变,增强数字图书馆的个性化服务水平并对知识服务的再拓展提供了巨大帮助,有助于推动数字资源的再开发和数据库的建设。医学高校数字图书馆知识发现系统在医学教育、医疗培训和医学科学研究方面发挥了重要作用,很多还为临床医生以及社会特定人群提供特殊的延伸服务,在用户需要时及时提供信息服务。医学高校数字图书馆以“互联网+”技术为基础构建的知识发现系统服务模式和运行机制,为用户提供了方便、快捷的知识服务,但其自身馆藏资源与用户信息也受到了各种安全威胁,因此数字图书馆应构建坚固的防御系统,随时避免自身的知识资料及用户个人信息等数据遭受泄露篡改,为用户提供更安全、更有效的数字信息服务。

“互联网+”时代,医疗信息化和教育信息化建设促使医学高校势必要过渡到智慧校园建设,而医学高校数字图书馆服务也必然要随着智慧校园的建设提高其智慧化的程度,具备安全数据处理能力的医学高校数字图书馆是其智慧化的重要体现。随着互联网技术的发展、社会性学习软件的普及应用、智能终端设备的普及、无线网络的覆盖、物联网及智慧校园的建设,未来医学高校数字图书馆服

务的发展状态应该是将不同用户群体在基础需求上所衍生出的个性化、差异性的行为信息融入到定制式服务作为发展和研究焦点,由此可见建立安全高效的知識发现系统将会成为医学高校数字图书馆智慧化发展体系中非常重要的部分,是保持医学高校图书馆繁荣的关键。

## 参考文献

- 1 IFLA World Library and Information Congress 80<sup>th</sup> IFLA General Conference and Assembly World [EB/OL]. [2017-03-21]. <https://www.ifla.org/past-wlic/2014/ifla80.html>.
- 2 The Hague Declaration on Knowledge Discovery in the Digital Age [EB/OL]. [2017-03-21]. <http://thehague-declaration.com/the-hague-declaration-on-knowledge-discovery-in-the-digital-age/>.
- 3 赵国荣. 高校图书馆数字资源整合探索与实践 [J]. 图书馆学研究, 2012, (20): 53-56.
- 4 张晓林. 走向知识服务: 寻求新世纪图书情报机构的生长点 [J]. 中国图书馆学报, 2000, (5): 148.
- 5 文丽. 医学图书馆嵌入式服务研究 [D]. 长春: 吉林大学, 2016.
- 6 李力. 医学图书馆嵌入式学科服务的案例分析与启示——以约翰霍普金斯大学 Welch 医学图书馆为例 [J]. 图书馆学研究, 2013, (17): 95-98.
- 7 Gijon-Correas JA, Andrade-Navarro MA, Fontaine IF. Alkermio: association of chemicals with biomedical topics by text and data mining [J]. Nucleic Acids Res, 2014, (42): 422-429.
- 8 Theodosiou TI, Vizirianakis IS, Angelis L, et al. MeSHy: mining unanticipated PubMed information using frequencies of occurrences and concurrences of MeSH terms [J]. J Biomed Inform, 2011, 44 (6): 919-926.