

图书馆系统的过去、现在与未来*

姜爱蓉

(清华大学图书馆, 北京 100084)

摘要: 阐述图书馆系统的三个发展阶段, 从20世纪80年代ILS系统的缘起和OPAC的演变, 到21世纪以来各种数字资源管理系统的发展和技术特点。重点阐述下一代图书馆系统的缘起、需求、关键要素、设计思路和技术特点。

关键词: 图书馆系统; OPAC界面; 资源整合; 发现系统; 图书馆服务平台; SOA架构; 云计算

中图分类号: G250.7

DOI: 10.3772/j.issn.1673-2286.2015.08.001

1 引言

进入21世纪, 图书馆的信息环境发生了巨大改变。因特网的普及使得信息交流冲破了传统的模式和界限, 大规模数字化对文献资源建设产生了深刻影响。在国际范围内, 大量图书文献回溯转变为比特, 原生数字资源不断涌现。美国ARL研究型大学图书馆超出60%的资源经费用来购置数字资源。数字资源日益成为提供服务的主体, 图书馆业务向数字化、网络化转变。现如今, 图书馆的主要矛盾已不再是资源匮乏与用户需求日益增长之间的矛盾, 而是数字资源的海量、异构、分布、多样性与用户便捷选择和准确获取之间的矛盾。新信息环境下, 图书馆系统在概念、内容和范畴上都发生了巨大变化。

2 图书馆集成系统的演变与新一代OPAC

20世纪60年代, 随着计算机技术的应用, 图书馆开始借助计算机处理纸本书刊的编目、流通、公共查

询等业务环节。最初只是分模块批量处理各自图书馆某方面业务, 且为单机脱机状态, 与其他图书馆的计算机没有联系。80年代初, 伴随机读目录出现, 单一功能的系统逐渐发展成为集成管理系统 ILS (Integrated Library System)。经过几十年发展, 集成管理系统的各个功能模块已相当成熟。传统ILS的理念与特点有: ①以本地馆藏管理为宗旨 (Local Collection Management); ②以书目控制为核心 (Bibliographic Control); ③以管理印刷型书刊为主体 (Managing Printed Materials); ④以系统内共享数据为目标 (Data Shared by ILS Components)。传统ILS主要致力于两个方面: 一是图书馆内部馆藏管理, 二是本机构内读者的管理。

20世纪90年代以来, 以数字图书馆为代表的信息技术发展迅速, 图书馆资源和信息环境开始发生转变, 以Google为代表的搜索引擎、以Web2.0为导向的各种网站受到用户的青睐, 提升了用户体验。新技术环境的变化给图书馆ILS系统带来压力, 用户希望图书馆系统能提供Google-like的检索以及与Web网站同样好的用户体验, 提供与外部系统的关联等。在技术发展和用户

* 本研究得到国家科技部项目“学习需求驱动下的数字出版资源定制投送系统及应用示范” (编号: 2013BAH47F00) 资助。

需求的驱动下,国内外许多图书馆针对ILS系统进行了拓展功能、丰富内容、优化界面、增强交互和增值服务等各种努力,目的是提升ILS的可用性和用户体验。这些针对ILS系统的变革主要包括:支持书刊的全面书目信息查找(包括封面、目次和摘要等);支持分面、主题云图等途径的检索和导航;支持相关度、使用率、用户行为等指标排序和推荐;检索结果的FRBR化呈现(按照书目层次关系显示);检索结果不理想时,给出其它查询选择(扩展检索词、检查拼写错误并提示、扩充检索范围等);增加用户评级、增加读书笔记和添加书评的功能;将用户需求直接链到在线书商、咨询馆员、馆际互借和文献传递(ILL/DDS)入口;提供图书所在书库和书架的位置图等。

最早对ILS系统进行变革的是美国北卡罗莱纳州立大学图书馆,2006年该馆发布了基于开源软件Endeca的新一代OPAC(如图1所示)。在随后的几年里,数百上千家图书馆或采用商业软件或采用开源软件建立新一代OPAC界面^[1]。

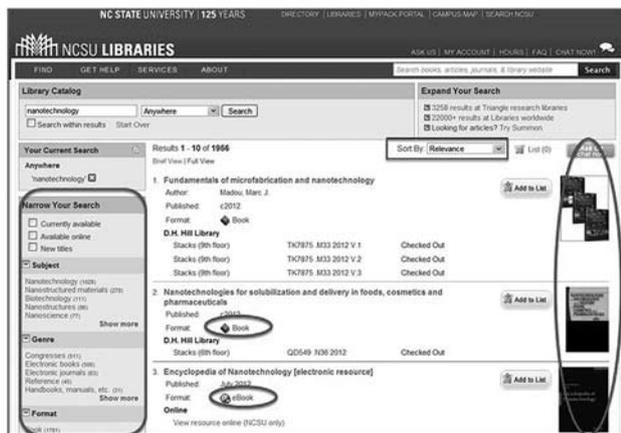


图1 美国北卡罗莱纳州立大学图书馆的OPAC界面

新一代的OPAC界面使得ILS系统摆脱单调、古板的形象,在分面导航浏览、灵活定制排序、可视化、交互性等方面大为改观,贴近了新技术和网络搜索引擎,开创了ILS系统应用的新局面。

3 图书馆集成系统的变革与出新

进入21世纪以来,受大规模数字化的影响,图书馆的资源结构发生了历史性变迁。清华大学1912年建立图书馆,经过103年的发展,实体馆藏逾470万件册,年度购置纸本图书逾7.8万种(约11万册)。自20世

纪90年代末期开始购置数字资源,经过近20年的发展,外购和自建的数据库逾510种,全文电子期刊逾6.7万种,电子图书逾870万册种,以及大量会议录、学位论文、标准、年鉴等(以上数据统计截止到2014年底)。传统图书馆主要是在书目层次对实体馆藏进行揭示和管理,数量级在几十万、几百万或者上千万,数字图书馆需要在文章层次(甚至更小粒度)对数字资源进行揭示和管理,数量级在上千万、上亿或者几十亿。揭示和管理这些海量数字资源对图书馆是一个严峻挑战,迫切需要得心应手的理想工具/系统。

20世纪90年代以来,全球范围内的数字资源在短时期内迅速增加。人们对数字资源的组织和揭示能力还远跟不上数字资源的增长速度。数字资源与传统的书刊文献不同,每种资源都有其物理和逻辑结构,依赖于各自的软件系统并借助于网络传播。数字资源带有与生俱来的异构性、分布性和资源描述的多样性。相对于实体馆藏,数字资源的揭示管理更为复杂。多年以来,图书馆一直在千方百计地寻求各种方法解决数字资源增长与整合揭示不足的矛盾,多途径探索组织和揭示数字资源的方法,期望能够应对资源不断变化的本质、满足用户不断提升的需求。例如,清华大学图书馆自1998年开始,先后建立了“电子资源/数据库导航系统”和“电子期刊导航系统”,2000年前后开始在“图书馆集成管理系统”中进行电子资源编目,2006年引进了“数字资源整合系统(MetaLib,联邦检索技术)”和“开放链接系统(SFX)”,2008年试用“电子资源管理系统(Verde)”,2011年引进“资源发现系统(Primo,元数据仓储预索引技术)”,2014年建立“电子图书导航系统(基于Primo)”。在探索实践数字资源揭示管理过程中,对数字资源的认识和理解也在不断深入。

“电子资源/数据库导航系统”局限在数据库名称层面上揭示资源,“电子期刊导航”和“电子图书导航系统”局限在文献的集合级(专著名称)和连续级(期刊名称)层面上揭示资源,对数字资源进行馆藏编目在揭示层面上与导航系统基本相同。这些层面上的揭示,无法满足用户快速定位到目标信息(文章的题名、作者及内容)的需求。用户需要更深层面、更细粒度、更小单元的资源揭示、查找和定位。“数字资源整合系统”和“资源发现系统”从分析级(文章题名、主题、文摘、作者等)进行统一检索和揭示,能够满足用户快速定位并获得文章信息的需求。在计算机检索能力(全文检索)已相当发达的今天,图书馆需要向用户提供查找深层

次信息的途径(目前是文章级,未来发展采用XML化和碎片化技术细化到微内容)。整合系统和发现系统的区别在于采用的技术手段不同,前者是依赖计算机的处理能力,将检索式分发到各个数字资源平台进行逐一检索,再将分别返回的检索结果进行合并显示。这种技术框架下,检索结果的返回速度受网络传输和原数据库平台响应的制约,影响了用户体验。后者是采用预先收集数字资源元数据并建立预索引,检索的实现和响应都在发现系统平台上集中进行,显著提高了检索效率和用户体验。

近几年,发现系统在用户需求和信息大环境的驱动下得到快速发展。一方面,数字资源增长迅速且价格昂贵,图书馆需要提升数字资源的可见度和易用性,提高使用率;另一方面,在资源数字化程度越来越高的大环境下,数字资源提供商越来越开放,他们愿意贡献和分享元数据,期待通过更多途径吸引更多用户使用其资源。在这样的背景下,资源发现系统应运而生。发现系统(Discovery System)的突出特点是快速、简洁、易用,有人称之为图书馆的Google。美国图书馆技术专家马歇尔·布利汀(Marshall Breeding)2013年在分析图书馆系统的市场时指出:发现服务继续在图书馆行业扮演主要的角色,对于有大量电子资源的图书馆,目前集中索引式的网络级发现服务已是“必须有(must-have)”的产品。自2008年/2009年发现系统问世,仅6-7年间,全球范围内的拥有发现系统用户逾万家,国内采用国外发现系统用户逾百家,采用超星发现系统用户逾300家。发现系统机构用户总数的快速增长足以说明图书馆和用户对数字资源统一发现的强烈需求。

在美国,越来越多的大学图书馆正在将发现系统作为“搜索学术信息的第一入口”,通过发现系统全面集成图书馆资源。这些图书馆通过API接口将本馆各种类型的资源仓储(或索引)与发现系统的元数据仓储汇集在一起,为用户提供全面的资源发现视图。在发现系统的技术框架上采用模块化组合,例如,有的大学采用开源的Blacklight做发现系统界面,配备ProQuest公司的Summon元数据仓储或者配备ExLibris公司的Primo Central元数据仓储。也有的大学采用ExLibris公司的Primo做发现系统界面,配备Summon元数据仓储。还有的大学采用开源的VuFind做发现系统界面,配备Primo Central元数据仓储。北卡罗莱纳州大学采用开源的QuickSearch做发现系统界面,配备Summon元数据仓储(如图2、图4所示)^[2]。该校发现系统的检索结

果按文献类型显示(如图3所示),条理清晰,用户可根据需要选择浏览。

发现系统的理念与特点主要有:①以全球学术信息整合为宗旨(Global Academy Information Integration);②以元数据控制为核心(Matadata Control);③以电子资源管理为主体(Managing

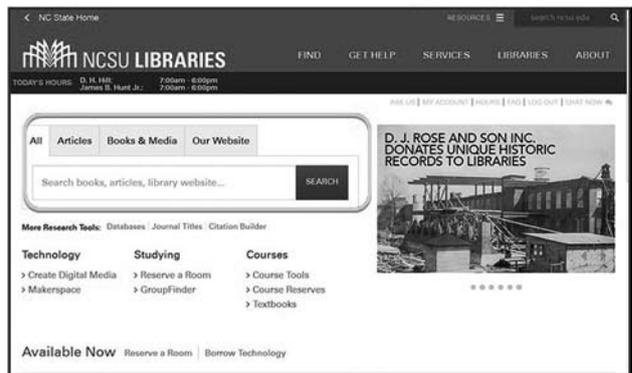


图2 北卡罗莱纳州大学图书馆主页

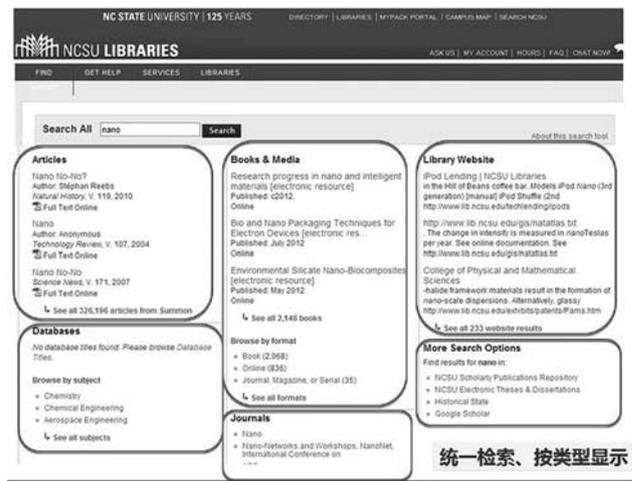


图3 北卡罗莱纳州大学图书馆发现系统结果显示页

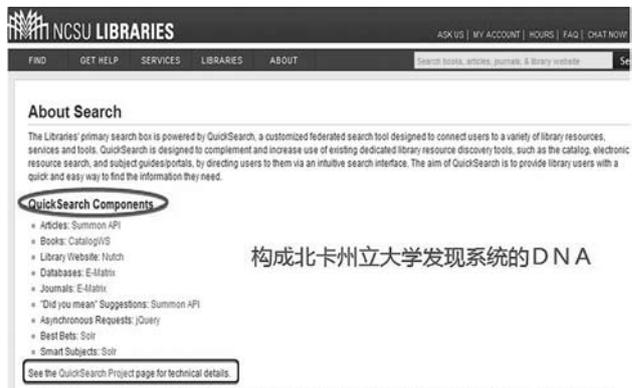


图4 北卡罗莱纳州大学图书馆发现系统的构成

Electronic Resources)；④以全球范围共享数据为目标 (Data Shared by Worldwide)。发现系统主要致力于三个方面：一是在全球层面上发现资源；二是与本地层面提供的目标对象 (全文内容) 建立链接保障；三是针对各种资源本身元数据进行集成、组织和揭示。

4 关于下一代图书馆系统

经历了ILS全面揭示管理纸本馆藏的过程，经历了多种数字图书馆系统各自独立、各见所长管理数字资源的阶段，图书馆期待着通过一个功能强大的系统来管理包括纸本和数字资源、拥有的和可用的资源以及所有图书馆赖以提供服务的资源。面对海量快速增长的各种资源，用户期待着通过一个直观、流畅的界面和高效准确的平台获取他们所需要的资源。由此，下一代图书馆系统提上日程，成为近几年图书馆领域的技术关注热点。

传统图书馆集成管理系统固化的图书馆业务流程，使其在数字图书馆环境中的可用性逐步减弱。传统ILS力所不能及的主要方面包括：①管理全媒体时代图书馆面对的所有馆藏类型；②管理与数字资源相关的获取、揭示、入库、整合、利用等整个生命周期的业务流程；③实现统一用户认证、门户建设与管理、资源整合与发现、统计分析 with 反馈优化等。现如今，许多图书馆的纸本馆藏和数字资源的揭示、组织和管理是分离在不同系统上由多个业务流程完成。在图书馆书目集成管理系统 (如INNOPAC) 上进行纸本资源的采编流管理和部分数字资源的编目，在电子资源管理系统 (如Verde) 上进行数字资源的采购管理，在数字资源整合检索系统 (如MetaLib) 上进行数据库的揭示和编目，在开放链接系统 (如SFX) 上进行电子期刊的揭示和管理，在资源发现平台 (Primo) 上提供整合检索和获取服务。图书馆的自建数字资源、数字资源的长期保存、用户认证访问等也分别建立在不同的系统平台上。这种系统各异、流程分离的局面对图书馆资源揭示的完整性和数据更新的及时性带来影响，迫切需要改变。图书馆期待在新格局的信息环境中，实现一体化的数据 (Unified Data)、一体化的工作流程 (Unified Workflows) 和一体化的管理 (Unified Management)，从而节省图书馆员的投入，降低人力成本，提高管理效率，提升服务质量。马歇尔·布利汀认为：“当前，对来自不同系统、不同格式的多类型资源的

揭示、集成和整合成为图书馆核心任务之一，是未来图书馆系统需要重点解决的问题”。

从图书馆业务流程角度分析，对下一代图书馆系统的需求包括：①支持纸本和数字资源的统一采访、编目、服务和统计分析；②支持全流程的数据复用、数据套录、数据深度融合；③支持数据之间的层级递归和关联关系；④支持多层面的导航和发现服务等。以电子期刊的馆藏编目为例：目前大多数图书馆是借助外部其他系统的数据，通过整理和规范化处理，批量导入图书馆ILS系统。用户检索命中某种期刊后需链接到外部期刊导航系统才能定位到年卷期，再链接至数据库平台才能浏览卷期目次和全文。期待新系统：能够提供期刊的全局知识库，图书馆员编目只需选择刊名、年期范围、载体类型和使用范围等；新系统支持建立期刊导航，支持直接浏览期刊的卷期和目次，支持直接链接到数据库平台浏览全文。针对期刊文献，融会贯通不同载体、不同文献级别、从传统系统仅仅揭示刊名卷期发展到新系统能够直接揭示目次文章信息并非易事，不仅需要庞大的期刊级和文章级知识库的支撑，而且需要建立期刊变迁历史之间的关联，需要建立期刊卷期和文章的层级递归和链接关系。

再以图书的采选为例：最近几年，图书馆的资源建设正在向“需求驱动的采购 (Demand-driven acquisition, DDA)”过渡。无论是纸本图书还是电子图书，图书馆在采购中都面对着数量庞大、选准选对十分不易的局面，用户参与图书馆的资源选购是一个科学合理的机制，正在得到国内外同行的广泛认同。有些图书馆正在尝试将书商提供的纸本图书征订书目导入ILS系统实现用户推荐，也有些数字资源商开始在电子书平台上提供按需采购模式，例如EBL、Ebrary等。这些现象都反映了图书馆的资源采购迫切需要一个强大的推荐系统支持。期待新系统，提供全球范围内书商 (资源商) 的征订信息，即全局性的资源采选知识库，方便用户查询和推荐图书 (资源)，方便图书馆学科馆员审核和采访馆员确定订单。有这样一个全球范围内书商 (资源商) 的书目 (资源) 信息知识库，省去了各个图书馆的采访馆员分别从四面八方获取征订信息的投入，同类型的图书馆或者服务于相同学科背景用户的图书馆，也可以相互参照和分享订购图书 (资源) 的经验。通过联机编目共享编目数据是图书馆领域的一大进步，期待联机采购共享图书 (资源) 征订信息也带来业务发展的进步。大数据时代，云计算模式下的新一代

图书馆系统,需要借助各层面的知识库,减轻图书馆员的工作,助力图书馆服务创新。

2012年以来,图书馆技术专家马歇尔·布利汀提出下一代图书馆系统正在向图书馆服务平台(Library Service Platforms,简称LSPs)过渡,这个观点的提出,反映了在图书馆技术环境的转变下,图书馆的业务流程和用户服务所需要的系统支持已远不是一个单独的系统所能胜任的。马歇尔·布利汀认为图书馆服务平台应具备五个特性,包括:①高度共享的数据模型(Highly Shared Data Models);②采用软件即服务的方式传递(Delivered through Software as a Service);③跨数据格式与媒体格式的统一工作流程(Unified Workflows across Formats and Media);④灵活的元数据管理(Flexible Metadata Management):包括机读格式(MARC)、都柏林核心元数据(Dublin Core)、视觉资源元数据(VRA)、元数据对象描述纲要(MODS)与在线信息交换标准(ONIX);⑤可扩展和互通的开放API(Open APIs for Extensibility and Interoperability)^[3]。

图书馆领域对下一代图书馆系统的期望值很高,新系统的关键要素包括:①采用面向服务的体系框架(SOA)重新设计系统,重构图书馆对各类资源管理的业务流程,而不是在传统ILS系统上改造;②实现所有类型资源元数据的统一管理,以全球知识库代替分散的本地资源库,支持多层面的资源发现和获取服务;③以软件即服务(SaaS)和云计算架构进行部署,通过API接口整合和嵌入机构的信息环境(课程、研究等),延伸图书馆服务;④从传统的“信息管理”过渡到“智能分析”。实时统计包括资源商、图书馆业务和用户使用等交易或交互数据,以及所有活动(包括点击流)的潜在记录,通过收集和分析数据、检测趋势,提供业务发展的决策建议^[4]。新一代图书馆服务平台的应用,将使现有的图书馆业务流程发生转变,这也将给图书馆的组织结构带来变革。

下一代图书馆系统采用面向服务的体系架构(SOA)(如图5所示)。在分布的系统环境中,SOA架构中的各种组件和子系统“松散”耦合,每个服务相对独立、自包含、可重用,由一个或者多个分布的系统所实现。SOA架构能使所有的资源和集合展现更灵活的工作流程,更便于与其它系统进行互操作,也更便于有效地进行系统管理(也就是在一个服务出现故障的时避免引起连锁反应),更好地控制系统内部的服务与服务

务之间的关系与优先等级。

云计算对下一代图书馆系统的部署产生着重大影响。下一代图书馆系统从逻辑上可划分为统一资源管理、统一资源发现和获取。侧重图书馆内部业务和数据管理、数据关联的统一资源管理(后端业务)完全采用云模式远程部署和使用;侧重用户服务的统一资源发现和获取(前端服务)可以部分云模式部署、部分本地部署。前端服务(用户)与后端业务(知识库和馆员业务)松耦合,前端系统为用户提供直观便捷的搜索界面,引导其快速发现所需资源。后端系统云模式部署,为图书馆提供协同工作方式,通过领域内和社区内的数据共享与合作服务,减少馆员投入成本,提高效率,聚焦更专业的核心服务。

目前国外开发商和图书馆合作,已推出和正在研发的下一代图书馆系统有:Ex Libris公司的Alma、Innovative Interface公司的Sierra、开源计划KualiOLE、OCLC的WorldShare Management Services(简称WMS)、Serials Solutions公司的Intota等^[5-7]。据不完全统计,国际范围内逾千家图书馆已经或正在迁移至下一代图书馆系统。至2015年初,Ex Libris公司Alma的用户数为420家,Innovative公司Sierra的用户为495家,OCLC的WMS的用户数为270家。我国图书馆领域自2013年开始关注下一代图书馆系统的发展,2014年以来进入调研、测试、比较、选型阶段。

5 结语

一直以来,图书馆的发展深受新技术和文献载体的影响。日新月异的新技术和不断演进的文献载体赋予图书馆新的使命,丰富图书馆的业务内容和服务形式,给图书馆带来了生机、挑战和机遇。现如今,随着数字资源日益增多,图书馆服务呈现多样化发展态势,

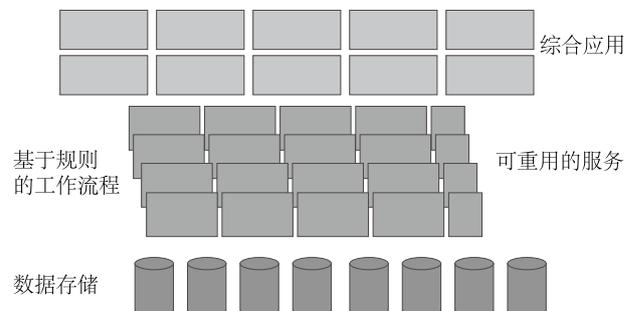


图5 面向服务的体系架构(SOA)

业务管理的复杂性加大。采用基于先进技术开发的下一代图书馆系统,实现对全部馆藏和可提供服务资源的统一管理,实现图书馆业务的协同工作,为用户提供更流畅、更及时的图书馆服务是大势所趋,也是新一轮挑战。下一代图书馆系统还在初始发展阶段,对关联数据的支持,对自建资源数字仓储的支持,对数据深度融合的支持等等都有待开发和完善。图书馆只有跟踪新系统的发展,保持对技术的敏感度,才能在新信息环境下形成新的服务能力和竞争力。

参考文献

- [1] Marshall Breeding. Library Technology Guides [EB/OL][2015-07-28]. <http://librarytechnology.org/discovery/>.
- [2] NCSU Libraries Homepage [EB/OL][2014-06-05]. <http://www.lib.ncsu.edu/>.
- [3] Marshall Breeding. Library Systems Report 2014 [EB/OL][2015-07-28]. <http://americanlibrariesmagazine.org/2014/04/15/library-systems-report-2014/>.
- [4] LibTechRFP. Unified library resource management specification [EB/OL][2015-07-28]. <https://libtechrfp.wikispaces.com/Unified+library+resource+management+specification>.
- [5] 殷红,刘炜. 新一代图书馆服务系统:功能评价与愿景展望[J].中国图书馆学报, 2013(5):26-33.
- [6] 包凌,赵以安. 国外下一代图书馆自动化系统的实践与发展趋势研究[J].图书馆学研究, 2013(9):58-65.
- [7] 陈武,王平,周虹. 下一代图书馆服务平台初探[J].大学图书馆学报, 2013(6):82-87.

作者简介

姜爱蓉,女,1955年生,清华大学图书馆研究馆员,研究方向:数字图书馆,E-mail: jiangar@lib.tsinghua.edu.cn。

Evolution and Development of Library System

JIANG AiRong

(Library, Tsinghua University Library, Beijing 100084, China)

Abstract: This paper introduces the three stages of the development of the library system, the origin of the ILS system and the evolution of OPAC Interface since 1980s, and development and technical characteristics of various types of digital resource management system since 21st century. This paper mainly introduces the origin, demand, key issue, design ideas and technical characteristics of the next generation of Library.

Keywords: Library Systems; OPAC Interface; Resource Integration; Discovery System; Library Services Platform; SOA Architecture; Cloud Computing

(收稿日期: 2015-08-10; 编辑: 雷雪)