

基于微服务架构的资源发现系统平台构建研究*

孙宇 周纲

摘要 基于微服务架构的第三代图书馆服务平台的研发、建设和实施,推动了图书馆资源发现系统技术的飞跃性发展。第三代图书馆服务平台具有全媒体管理、全流程管理和全网域发现等特点;微服务架构通过统一的互操作标准,实现基于总线管理的多个不同功能模块的敏捷开发和更新;基于微服务架构的资源发现系统可以更灵活地实现异构系统、图书馆知识库等多平台资源的发现和获取。本文设计构建了新一代图书馆资源发现系统的架构,系统采用数据层、支撑层、接口层、应用层和门户层的五层架构设计,实现基于 Codex 检索的图书馆知识库和开源知识库的资源统一发现获取。文章同时介绍了本地化实现和其他功能扩展,包括统一用户管理、界面设计,以及基于关联数据的可视化展示与分析等。基于各个独立知识库的资源发现系统的构建,能更好更灵活地整合图书馆的电子资源、馆藏资源以及开放资源,通过构建基于微服务架构的各类 APP 实现多样化的用户服务。图 3。参考文献 17。

关键词 图书馆服务平台 FOLIO 资源发现系统 知识库

分类号 G250.7

The Construction of Resource Discovery System Platform Based on Microservices Architecture

SUN Yu & ZHOU Gang

ABSTRACT

The third generation library services platform (LSP) is a service platform that integrates all types of resources (including paper / electronic / digital, purchase / self built / shared / Open Access) available to all users under the guidance of user demand and the design concept of "user centered", covering the functions of digital resource management and resource discovery. The third generation LSP has three necessary functions: all media management, all process management and all domain discovery. The resource discovery system based on the third generation LSP can flexibly achieve the harvesting and acquisition of all kinds of resources. Combined with the unified metadata retrieval and federated retrieval, and the microservices system can realize the disclosure and discovery of multi system and multi platform resources.

The purpose of this paper is to explore the structure and composition of the new generation of library

* 本文系国家自然科学基金一般项目“面向数字人文研究的图书馆开放数据体系构建与服务模式设计研究”(编号:18BTQ027)和国家自然科学基金重大项目“面向大数据的数字图书馆移动视觉搜索机制及应用(模式)研究”(编号:15ZDB126)的研究成果之一。(This article is an outcome of the project "Research on Library Open Data System Construction and Service Model Design for Digital Humanities Research"(No. 18BTQ027) and the major project "Research on Mobile Visual Search Mechanism and Application of Digital Library for Big Data"(No. 15ZDB126) supported by National Social Science Foundation of China.)

通信作者:孙宇,Email:ysun@libnet.sh.cn,ORCID:0000-0001-9409-990X(Correspondence should be addressed to SUN Yu,Email:ysun@libnet.sh.cn,ORCID:0000-0001-9409-990X)

resource discovery system. This paper first introduces the evolution and development of resource discovery system. Next, the paper introduces the characteristics of the third generation LSP and the advantages of microservices architecture. The third generation LSP is represented by FOLIO (the Future of Library Is Open). On this basis, a resource discovery system platform based on FOLIO microservices architecture is constructed. The system is divided into five layers: data layer, support layer, interface layer, application layer and portal layer. Through the core microservices domain Codex in the FOLIO resource management system, cross domain data interaction is realized. Through the knowledge base management module of FOLIO, the construction of metadata warehouse knowledge base is realized, and the unified management of all kinds of digital resource databases, self built resource databases, collection characteristic databases and institutional databases is realized.

This paper also introduces the localized implementation of the service platform of the new generation resource discovery system, including academic resource discovery, collection resource discovery based on Codex query, system using PostgreSQL database, data storage in Jsonb format, and the construction of unified resource retrieval based on knowledge base. Next, the paper introduces other function extensions of the new generation resource discovery system service platform, including unified user management, multi window format interface design, visual display and analysis, etc.

Finally, the development of the system is prospected. The system needs to expand and improve its functions through various flexible APPs/API interfaces. The construction of resource discovery system based on each independent knowledge base will better and more flexibly integrate the electronic resources, collection resources and open resources of the library, and realize diversified user services through the construction of various types of APPs based on microservices architecture. By building a resource discovery platform based on microservices architecture, based on various kinds of knowledge base at the bottom, and through the system bus service and OAI-PMH API, it can flexibly realize the unified retrieval and discovery service of various resources in the collection of multi platforms and multi systems. 3 figs. 17 refs.

KEY WORDS

Library service platform. FOLIO. Resource discovery system. Knowledge base.

0 引言:资源发现系统的演进

2009年,基于全网域的资源发现服务(Web-Scale Discovery Services)系统引入图书馆^[1],该系统基于海量的元数据集,用来揭示馆藏数字资源,实现自建和外购资源的统一发现和获取。目前市场上影响较大的资源发现系统(Resource Discovery System, RDS)产品有EBSCO Discovery Service (EDS)、Primo、Summon、Worldcat Local、Enterprise、VuFind^[2]、超星发现、维普智立方等,这些产品具有检索界面易用,基

于权重的相关性排序,可以根据特定类别、著者或日期范围限定结果的分面^[3]等优势。

资源发现系统的研发以替代联邦检索为目的。联邦检索在资源检索时能够实现多个库的同时跨库检索,但主要问题是缺少结果的相关性排序,检索结果返回需要较长的时间,以及弱化了数据库本地接口组件提供的搜索功能等。资源发现系统从出版商、集成商、图书馆等处获取海量文章或文章元数据,对元数据进行抽取、映射、导入,能够形成一致的底层知识库,避免了不同资源不同数据结构的检索障碍。

资源发现系统经过多年的应用实践,其不

足也正在逐渐显露,如元数据的授权问题、资源的可获取性不足、数据库资源的远程访问受限,以及馆藏资源揭示的同步更新滞后等问题,影响了用户的使用和体验。基于元数据索引的发现系统的构建,解决了统一的跨库检索问题,推动了图书馆技术的发展,但同时其系统架构限制了资源的有效发现和揭示服务。

近几年,资源发现系统的应用逐渐向检索结果的多窗格布局模式发展,系统在发现服务平台的基础上,通过 APIs(Application Programming Interfaces)接口服务,实现了同时检索展示馆藏目录、数据库资源、自建资源等内容,比较灵活地实现了馆藏所有数字资源的统一检索发现服务^[4]。

1 基于微服务架构的新一代资源发现系统平台

1.1 基于微服务架构的第三代图书馆服务平台

资源发现系统应用以图书馆服务平台为基础,构建在图书馆服务平台之上。传统图书馆的资源管理依赖图书馆集成管理系统(Integrated Library System, ILS)完成对纸质图书与连续出版物的管理。但在数字化的大环境下,图书馆需要管理的数字化资源种类不断丰富,ILS 系统已无法适应多类型资源管理的需求,因而出现了 ILS 系统与电子资源管理系统、机构知识库系统等多平台并存但独立运行的现状,数据孤岛现象制约了图书馆的服务效率。第三代图书馆服务平台(Library Services Platform, LSP)是在用户需求牵引下,采用“以用户为中心”的设计理念,整合所有用户可获取的各类型资源(包括纸质/电子/数字化和购买/自建/共享/开放获取),涵盖了数字资源管理和资源发现功能的服务平台。第三代图书馆服务平台具有全媒体管理、全流程管理和全网域发现等三个必备功能。基于第三代图书馆服务平台的资源发现系统可以灵活地实现各类资源的收割、获取,结合元数据统一检索和联邦检索,架构微服务体系实现多系统、多平台资源的揭示和发现。

第三代图书馆服务平台以 FOLIO(The Future of Library Is Open)系统为代表。FOLIO 是 2016 年兴起的新型开放获取的学术交流合作项目,其通过建立支持区域协作的数据共建共享和云服务架构,支持各类基于应用程序接口(APIs)的应用接入,采用微服务架构形成统一的互操作标准和应用市场,是由用户、图书馆和厂商共同进行模块化功能组装、开发与维护的平台^[5]。FOLIO 图书馆服务平台的建设目标是创建开发者可以协同开发的社区,降低系统开发的成本,通过更多的机构参与来改进产品,提供各种灵活的组件满足图书馆不同的需求选择^[6]。

FOLIO 采用微服务架构设计,具备敏捷开发的软件更新模式,项目的任何部分都不会因为太大而无法替代,即便有遗留问题也不会影响未来的开发。图书馆可以选择搭建自己的 FOLIO 平台,使用 FOLIO 提供的各种免费的 APP,或者采用第三方提供的收费的 FOLIO 平台和 APP。作为开放的图书馆服务平台,FOLIO 采用 RAML 标准,基于 Raml/RMB 生成接口定义实施开发,任何符合其规范和标准的接口 APP 都可以兼容 FOLIO。

微服务可以看作是一个“黑盒子”,各系统模块、存储模块、UI 模块等各类 APP 挂载在基于 Okapi 的 FOLIO 总线上。每一个微服务模块都可以有自己的存储,所有的微服务都通过 API 接口实现数据的交换。微服务将一个应用程序划分为很多小的服务模块,服务进程独立运行,各服务之间又相互协调和配合^[7]。微服务架构涉及的概念包括配置管理(Config Management)、服务发现以及负载均衡(Service Discovery & Load Balancing)、恢复与容错(Resilience & Fault Tolerance)、API 管理(API Management)、安全认证与鉴权(Service Security)、统一日志分析(Centralized Logging)、统一指标监控(Centralized Metric)、分布式服务跟踪(Distributed Tracing)、调度与部署(Scheduling & Deployment),以及自动缩放与自我修复(Auto Scaling & Self-Healing)等。在 FOLIO 中,Okapi 服务支持图书馆各类

服务的 APIs,例如 OAI-PMH API、采访/订单 API、读者服务 API、ERM API、资源发现 API、馆藏创建 API、馆藏借出/归还 API、馆藏查询 API 等^[8]。

1.2 基于 FOLIO 的新一代资源发现系统平台架构

基于图书馆微服务架构的资源发现系统平台架构如图 1 所示。

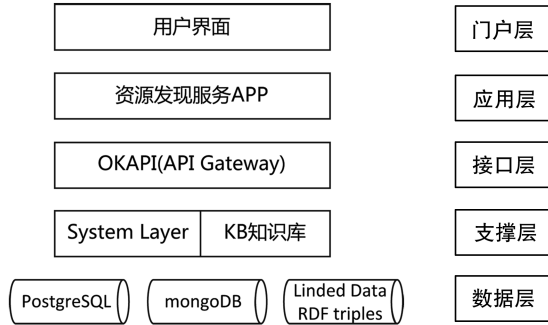


图 1 基于图书馆微服务架构的资源发现系统平台架构

系统的逻辑和层次结构功能详述如下。

(1) 数据层:底层存储馆藏元数据,以及通过授权、收割等方式获取的各类知识库元数据。事务性数据一般存储在开源分布式关系型数据库 PostgreSQL 中,书目馆藏数据采用关联数据格式存储在非关系型数据库 mongoDB 中^[9],并且支持 RDF 关联数据的存储。目前 FOLIO 系统内置 GOKb 知识库和 EBSCO 知识库。

(2) 支撑层:即系统层,实现各种业务逻辑。实现馆藏资源、电子资源、知识库(Knowledge base, KB)等各类元数据的映射、管理和整合发现,实现底层跨库检索,以及提供知识库用户认证和权限管理等。

(3) 接口层:通过系统总线,实现各类发现系统 API 接口的接入,如 EDS API,按照统一的 API 规范进行开发、接入和管理。

(4) 应用层:即发现层,接入各类发现系统 APP。基于微服务总线架构的 APP 模块,可以是一个完整的 APP,包含存储模块、系统模块和 UI 模块;可以是一个服务 APP,包含存储模块和系统模块;或是一个事务逻辑 APP,即仅包含系统模块;也可以是一个只包含 UI 模块的 UI APP。在这之上,发现系统可以自由架构需要的各类服务。

(5) 门户层:即用户展现层,FOLIO 的界面设计工具是 Stripes。作为用户使用的 UI Toolkit,资源发现系统界面开发可使用 VuFind、Blacklight、EDS UI 等。

通过 FOLIO 微服务的总线架构,新一代资源发现系统设计要能够支持异构数据系统的统一揭示和发现,实现异构数据的共享和透明访问。图书馆知识库包含了外购的资源、自建的资源、馆藏书目资源等任何一种内容提供形式,这些资源数据的来源和格式可以都不相同。这些数据库具有自治性,在微服务总线架构下,实现数据统一发现时,每个数据库系统都能保持自身的完整性和安全性,以及自己的应用特性。对于用户来说,所有的资源整合都是透明的,在保证可访问性的前提下,用户使用的是一个整合的资源库。

1.3 Codex

FOLIO 在其资源管理体系中定义了一个核心的微服务域——Codex^[10]。Codex 是一种标准化和虚拟化层,它允许 FOLIO 集成各种资源的元数据,无论其格式、编码或存储位置如何。FOLIO 通过 Codex 中介模块进行跨域的数据交互,Codex 可以充当不同模块之间的协调者。每个模块只需与中介模块集成,即可实现对其他模块的数据

需求,但又无需了解其他模块的业务。

Codex 是 FOLIO 的资源中心,它可以集中管理权威词表、受控词表。FOLIO 其他资源管理模块可以将其数据记录映射成 Codex 域中对应的记录,例如 Inventory 典藏、eHoldings 电子资源或 MARCcat 编目等模块,由此其他模块都可以平等地使用这些数据服务。FOLIO 基于 Codex 为资源提供统一的搜索功能,不管负责管理它的特定模块是什么,都可以提供一个单一的入口点来定位整个 FOLIO 系统中的任何资源。

1.4 知识库管理

知识库管理是通过建设元数据仓储知识库,完成图书馆各类数字资源数据库建设,包括对自建资源库、馆藏特色库、机构库的统一管理。底层元数据仓储涵盖期刊论文、学位论文、会议论文、科技报告、专利、标准、政策法规、科技成果、图书专著、视频等各种类型资源。通过对元数据仓储知识库的加工整理,图书馆能够根据需求对学科、主题、人物、机构、基金等知识获取进行数据的深度标引和质量检查,以作为海量科技文献统一检索和知识挖掘分析服务的数据基础。知识库是现代图书馆工具和电子资源系统的核心。

FOLIO 的知识库管理,从系统本身来说,包含了 GOKb(开源)和 EBSCO KB,以及馆藏(包括电子/纸质书刊和光盘等非书出版物)知识库,属于 Title 级别的出版物集合。目前,FOLIO 自带的知识库存储在 eholding 模块中,需要有数据源系统才能够运行,数据源通过接口方式提供服务,EBSCO KB 通过 HoldingsIQ 接口提供数据服务^[11]。其他各类图书馆知识库,如馆藏资源知识库、电子资源知识库、各类自建资源数据库等,都可以通过知识库模块进行管理。

FOLIO 在 Agreement APP 中管理 GOKb 协议和电子资源。GOKb^[12](Global Open Knowledgebase,全球开放知识库)是 Kuali OLE 和 JISC (Joint Information Systems Committee,英国联合信

息系统委员会)合作开发的一个国际开放数据存储库,是一个可由图书馆获取和管理的与电子资源相关的开放数据存储库。GOKb 提供的所有数据都是根据 CC0 1.0 公共领域开放数据专用贡献条款提供的。目前 ONIX-PL 是 GOKb 中许可证的唯一导出格式。2013 年安德鲁·W·梅隆基金会授予 NISO (National Information Standards Organization,美国国家信息标准组织)在 ONIX 出版物许可证(ONIX-PL)格式里支持电子资源集合模板许可证的编码方案,存入 GOKb 和 KB+知识库,免费分发给图书馆、出版社和图书馆系统的社区,允许许可电子内容的图书馆将模板许可证导入它们自己的电子资源管理系统,以便进一步进行本地定制和实现^[13]。NISO 通过制定 KBART(Knowledge Bases and Related Tools,知识库和相关工具)最佳实践,能够为 OpenURL 链接解析器提供数据,管理发现系统元数据,确保知识库的完整性和功能性^[14]。GOKb 还可以作为一个平台,探索开放知识库数据如何为更广泛的学术社区基础设施做出贡献,特别是围绕开放访问(Open Access, OA)开展相应的工作。图书馆自有知识库可以导入 FOLIO 的 Agreement APP 后使用。

FOLIO 在 eHoldings APP 中管理 EBSCO KB 协议和电子资源。通过 EBSCO KB API 访问 EBSCO 的 HoldingsIQ。EBSCO 知识库是可用的最大的已知标题数据池,允许图书馆管理它们的全文库。在多个应用程序中使用一个权威知识库可以减少重复和错误。多个应用程序管理来自不同数据源的 holdings 数据需要手工输入和重复的工作,HoldingsIQ 自动化了这个过程,将 holdings 数据集成到图书馆已使用的工具上^[15]。

2 新一代资源发现系统服务平台的本地化实现

由于数据来源的显著区别,以及获取途径的大相径庭,新一代资源发现系统在设计实现时,需要分别考虑外文学术资源、中文学术资

源,以及本地书目数据的发现和获取。图书馆利用 FOLIO 开放、多元的特性,可以在本地搭建多样化、个性化的资源发现服务平台。

2.1 学术资源发现

学术资源发现可使用 EBSCO SaaS 服务,通过 EDS API 接入,并进行发现系统功能优化;也可以使用基于 Codex 的 EBSCO KB 的学术资源统一发现系统。

FOLIO eHoldings APP 通过 EBSCO KB API 方式访问 EBSCO 的 HoldingsIQ,这对现有商业化学术资源和电子厂商具有借鉴意义,例如超星正在研发检索 API,探索实现 FOLIO APP 检索超星资源的可行性。

2.2 馆藏资源发现

相比目前资源发现系统或者公共查询目录 OPAC, FOLIO 的资源发现系统在架构上有其先天优势。符合 FOLIO 框架的 APP 中的资源无

需收割,可采用 Codex 单一的资源检索入口。首先,数据来源不一定局限在馆藏资源,而是 FOLIO 平台中所有 APP 管理的资源,例如图书采购信息、电子阅读设备信息、读者活动信息等元数据都可以映射到 Codex 数据域,都可以通过 Codex 检索到。其次,检索的方式相对统一,都利用 FOLIO 的微服务架构和统一格式的 API,未来在资源发现系统中很容易增加新的 FOLIO APP。

发现系统要代替目前的 OPAC 馆藏资源检索系统,将馆藏书目数据的管理作为底层知识库。实现 FOLIO 资源管理模块中的数据抽象层是 Codex 标准化元数据集,这些元数据来自于典藏(Inventory)、电子资源/纸本资源管理知识库等模块,元数据对象结构使用的是类似 BIBFRAME 的关联数据结构,包含 work-instance-item/holding 三层模型。基于 Codex 的查询功能,可实现知识库中馆藏电子资源、纸本资源的统一检索。基于 Codex 查询的系统架构如图 2 所示。

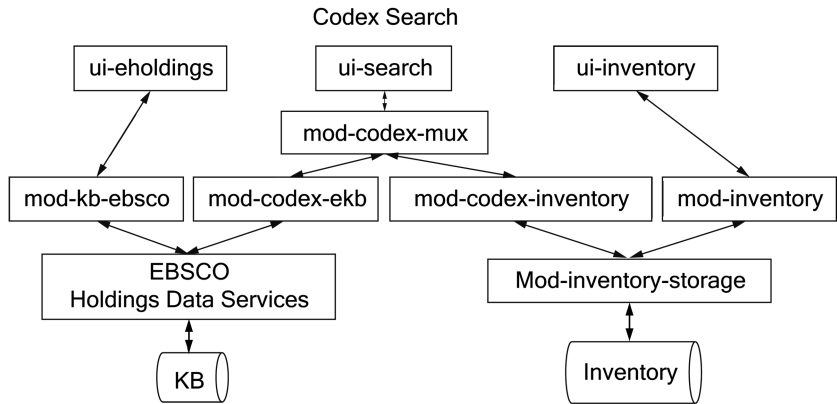


图 2 基于 Codex 的查询功能的设计实现

基于 Codex 的查询功能可实现从目前书目和馆藏元数据格式到关联书目数据格式的映射和转换,并和已有的资源发现系统功能进行整合。馆藏资源揭示需要设计实现馆藏书目数据到关联数据的映射,以及 BF/RDF 数据到 FOLIO-Inventory 的映射导入,最终通过 Codex Search 实现书目数据的检索发现。另外,还需要实现

资源发现功能中书目数据和馆藏信息等元数据的采集、清洗、转换和整合。实现从书目数据 MARC 到 Inventory,以及 BIBFRAME/RDF 数据映射导入等两个途径来源的元数据的采集,并对不完整数据、错误数据、重复数据进行清洗。通过建立一种 BF 的中间格式,来提高系统数据导入的效率,系统内部数据以 Jsonb 格式进

行存储,方便查询功能的实现。

FOLIO 通过 Codex 来集中管理数据,每个 APP 中的数据已映射到 Codex 元数据。传统资源检索系统的数据处理功能会前置到 FOLIO 某个 APP 的数据导入过程中。FOLIO 原生的数据导入(Data import)模块,尚不能定制各种数据源,也不能自定义数据映射关系。在此考虑增加(中间)元数据格式,将不同数据格式按照预定的规则,映射到一个统一的中间格式,再将数据导入到 FOLIO APP。采用统一的(中间)元数据格式,会使数据导入功能相对稳定。

传统的图书馆集成管理系统或者资源检索系统的数据格式有限,例如 MARC、DC、Excel 等。FOLIO 资源检索系统会接受更多的数据格

式,将数据的映射、处理和导入放在一个数据处理步骤中会使功能有局限性,为此 FOLIO APP 的数据导入建议分成两个步骤:(中间)元数据格式的数据映射和处理,数据导入。原始数据到(中间)元数据格式的处理可以屏蔽原始数据的多样性和复杂性,处理过程中进行数据的清洗、转换和整合,支持从目前书目数据 MARC、关联书目数据、BIBFRAME/RDF 等数据映射导入到(中间)元数据格式,并对不完整数据、错误数据、重复数据进行清洗。

FOLIO APP 使用 PostgreSQL 数据库,以 Jsonb 格式存储,当 APP 升级或元数据有变化时,不需要修改表结构。图 3 显示了导入到 FOLIO Inventory APP 的馆藏数据样例。

```
{
  "id": "8c4438c0-b064-4af1-ace2-6c18a657a8f8",
  "hrid": "in00001007",
  "notes": [],
  "title": "实木家具",
  "series": [
    "家具制造业生产管理全方案"
  ],
  "source": "BIBFRAME",
  "editions": [],
  "metadata": {
    "createdDate": "2019-04-15T07:58:20.485+0000",
    "updatedDate": "2019-04-15T07:58:20.485+0000",
    "createdByUserId": "0fe169ae-d782-5a65-84a0-9cea967965a2",
    "updatedByUserId": "0fe169ae-d782-5a65-84a0-9cea967965a2"
  },
  "statusId": "批量上传",
  "subjects": [
    "家具--制造业--工业企业管理"
  ],
  "languages": [
    "chi"
  ],
  "indexTitle": "实木家具",
  "identifiers": [
    {
      "value": "7-81079-644-5",
      "identifierTypeId": "8261054f-be78-422d-bd51-4ed9f33c3422"
    }
  ]
}
```

```
    }  
  ],  
  "publication": [  
    {  
      "role": "制作者",  
      "place": "<http://data.library.sh.cn/entity/place/hdg3vd6gkt74djx",  
      "publisher": ".:B9ef6250bd06dfc031345336fb69b2155",  
      "dateOfPublication": "2006"  
    }  
  ],  
  "contributors": [  
    {  
      "name": "<http://data.library.sh.cn/entity/person/8f4503p0zd2qs",  
      "contributorTypeText": "著",  
      "contributorNameTypeid": "2b94c631-fca9-4892-a730-03ee529"  
    }  
  ],  
  "staffSuppress": false,  
  "instanceTypeid": "6312d172-f0cf-40f6-b27d-9fa8feaf332f",  
  "classifications": [  
    {  
      "classificationNumber": "F407.886",  
      "classificationTypeid": "fb12264c-ff3b-47e0-8e09-b0aa074361f1"  
    }  
  ],  
  "electronicAccess": [],  
  "holdingsRecords2": [],  
  "modeOfIssuanceid": "9d18a02f-5897-4c31-9106-c9abb5c7ae8b",  
  "publicationRange": [],  
  "alternativeTitles": [],  
  "discoverySuppress": false,  
  "instanceFormatids": [  
    "a6facd90-0b47-4b9e-840c-64c03ec9ef90"  
  ],  
  "statusUpdatedDate": "2019-04-15",  
  "statisticalCodeids": [],  
  "physicalDescriptions": [  
    "表",  
    "474页",  
    "26cm"  
  ],  
  "publicationFrequency": []  
}
```

图3 以 Jsonb 格式存储的馆藏数据

2.3 基于知识库的资源统一检索

FOLIO 系统基于对知识库的管理,可以扩充到馆藏所有数字资源,通过 Codex 可以集成知识库和元数据仓储。知识库可以是任何一种内容提供形式,本地的典藏库、元数据仓储、数字人文数据库等都可以作为知识库进行统一管理,并通过微服务架构实现统一检索和发现。电子资源通过 FOLIO 标准的知识库实现书刊/出版物和本馆纸本资源的对应和关联,最终实现电子资源和纸本资源的统一揭示和发现。

底层 MOD-OAI-PMH 模块通过 OAI-PMH 2.0 开放文献元数据收割协议和 RESTful API,可以实现不同格式元数据的收割和存储^[16]。

3 新一代资源发现系统服务平台其他功能扩展

3.1 统一用户管理

新一代资源发现系统能够实现和图书馆本地用户认证系统的对接,如 OAuth2.0 或 LDAP 等,实现统一用户管理。对于远程用户(非到馆用户)的访问,发现系统通过本地代理服务器,结合 OpenAthens、Shibboleth 等基于用户 ID 认证的远程用户认证解决方案,代替目前的 VPN 登录方式。传统 VPN 远程登录方式受到终端、系统环境和资源数据库技术的约束,使得资源可访问性和获取性受到限制。另外一种基于“URL 重写”技术的代理认证方式,需要针对每个数据库分别配置转换地址,降低了资源的访问效率。新的认证方式的使用,能够解决目前各种移动终端资源的远程访问获取问题,用户只需要登录账号,就可以实现资源的远程获取。

系统对接的方式则可以根据应用场景,采用 GET/POST 请求方式设计实现接口功能。基于本地的代理服务器可以实现用户认证、使用行为分析,以及分组和权限控制等。

3.2 图书馆服务平台界面设计

图书馆可将资源发现系统作为图书馆的主

入口,帮助用户发现并使用图书馆资源;也可将资源发现系统 API 嵌入到第三方前端发现界面,实现无缝整合。可使用的界面设计工具有 Drupal、Vufind、Blacklight、Koha,以及图书馆网站集成解决方案 Stacks 等。随着移动设备的普及,支持各类移动终端是界面设计的核心要素。

界面设计可采用近年来“Bento 盒子”的设计理念,结合联邦检索和资源统一检索技术,检索结果采用多窗格布局模式,通过 APIs 接口服务最大限度地满足读者资源检索的需求^[17]。

3.3 可视化展示与分析

由于发现系统采用 Codex 技术实现数据的存储管理和统一发现,因此可设计实现基于 RDF 格式的数据可视化展现和数据分析。馆藏数据结构可采用关联数据技术进行存储,再加上底层知识库的集成,发现系统拥有灵活的数据分析和展现功能。发现系统可将知识图谱技术运用到内部数据聚合、外部数据采集、自动标引、数据挖掘分析、数据成果展示的流程中;通过知识图谱技术对各类元数据仓储进行深度挖掘,采用知识图谱分析算法,提升已有数据挖掘和机器学习方法的智能性,帮助用户识别更多有价值的知识,发现未知的事实或潜在关系;通过数据关联规则,进行数据关系描述,将数据知识以图的方式进行关联,通过知识图谱形式进行分析展示。

此外,发现系统设计并实现了方便的统计分析功能,可按用户检索词和所阅读文献(篇名、作者、文献来源、关键词、摘要等元数据)进行统计,分析获取用户的搜索和阅读行为,并按用户和数据库两个层面统计文献利用情况,包括资源的搜索、访问、浏览、阅读、下载数量等。

4 总结

改进资源发现服务,需要实现馆藏书目查询系统功能的升级,提高资源检索系统的稳定性,并提供灵活的页面、主题、检索和资源的设

置功能,以及资源整合、分面检索、导航等功能,提高读者的检索体验。资源发现系统在已有服务功能的基础上,优化检索结果展示,实现检索结果集的定制和纸电资源的统一揭示,提供标准 API 接口进行功能扩展,完善统计分析功能等。发现系统除了要实现为广大普通读者揭示发现资源的功能外,还要为高端用户群体提供专业的服务,如学科定制、个性化服务等功能。

未来系统架构需要通过各类灵活的 APPs/API 接口,实现功能的扩展和完善:通过构建各个独立知识库的资源发现系统,将更好更灵活地整合图书馆的电子资源、馆藏资源以及开放资源;通过构建基于微服务架构上的各类 APP

实现多样化的用户服务;通过构建基于微服务架构的资源发现平台,基于底层各类知识库,并通过系统总线服务以及 OAI-PMH API,能够比较灵活地实现多平台、多系统的馆藏各类资源的统一检索发现服务。

智慧图书馆的发展架构于资源和服务之上,这就需要更加灵活、可扩展的系统架构。构建微服务架构之上的新一代资源发现系统服务平台更加灵活,开发更加敏捷,通过实现全网域的资源发现,能够将图书馆服务不断扩展和延伸,不论是在资源的全面覆盖,还是在功能完善的各类服务方面,都能更好地推动智慧图书馆的发展。

参考文献

- [1] Gross J, Sheridan L. Web scale discovery: the user experience[J]. *New Library World*, 2011, 112(5/6): 236-247.
- [2] Burke J J, Tumbleson B E. Search systems and finding tools[J]. *Library Technology Reports*, 2016, 52(2) :17-22, 2.
- [3] Library Technology Guides. Major discovery products[EB/OL]. [2019-07-16]. <https://librarytechnology.org/discovery/>.
- [4] 刘伟成, Mischo W H, Schlembach M. 一种基于 Bento 模式的联邦检索系统[J]. *图书情报工作*, 2017, 61(19):113-121. (Liu Weicheng, Mischo W H, Schlembach M. A hybrid Bento-style federated search system [J]. *Library and Information Service*, 2017, 61(19):113-121.)
- [5] 谢蓉, 刘炜, 朱雯晶. 第三代图书馆服务平台:新需求与新突破[J]. *中国图书馆学报*, 2019, 45(3) :25-37. (Xie Rong, Liu Wei, Zhu Wenjing. The third generation library services platform: new requirements and new breakthroughs[J]. *Journal of Library Science in China*, 2019, 45(3) :25-37.)
- [6] FOLIO. About FOLIO[EB/OL]. [2019-07-30]. <https://www.folio.org/about/>.
- [7] 肖铮, 林俊伟. 用微服务构架下一代图书馆服务平台——以 FOLIO 为例[J]. *图书馆杂志*, 2018(11):63-69. (Xiao Zheng, Lin Junwei. Building the next generation library service platform with microservices architecture: taking FOLIO as an example[J]. *Library Journal*, 2018(11):63-69.)
- [8] FOLIO. Software versions[EB/OL]. [2019-07-31]. <https://folio-clover.aws.indexdata.com/settings/about>.
- [9] 肖铮, 林俊伟, 陈丽娟. 下一代图书馆开放服务平台 FOLIO 初探[J]. *图书馆学研究*, 2018(15):34-38, 63. (Xiao Zheng, Lin Junwei, Chen Lijuan. The research of FOLIO as a next generation library service platform[J]. *Research on Library Science*, 2018(15):34-38, 63.)
- [10] Bureau V. The Codex vision[EB/OL]. [2019-08-09]. <https://wiki.folio.org/display/PLATFORM/The+Codex+Vision>.

-
- [11] EBSCO API. Get started with our knowledge services[EB/OL]. [2019-07-03]. <https://developer.ebsco.com/gettingstarted>.
- [12] GOKb. About GOKb[EB/OL]. [2019-08-09]. <https://gokb.org/about-gokb/>.
- [13] NISO. ONIX-PL encoding project[EB/OL]. [2019-08-08]. <https://www.niso.org/standards-committees/onix-pl-encoding-project>.
- [14] NISO. KBART frequently asked questions[EB/OL]. [2019-08-08]. <https://www.niso.org/standards-committees/kbart/kbart-frequently-asked-questions>.
- [15] EBSCO. HoldingsIQ[EB/OL]. [2019-05-18]. <https://www.ebsco.com/products/holdingsiq>.
- [16] GitHub. Mod-oai-pmh[EB/OL]. [2019-09-03]. <https://github.com/folio-org/mod-oai-pmh>.
- [17] 孙宇. 公共图书馆资源发现系统发展探讨——以上海图书馆为例[J]. 新世纪图书馆, 2019(4): 75-79, 97. (Sun Yu. Discussion on the development of resource discovery system in public library; taking Shanghai Library as an example[J]. New Century Library, 2019(4): 75-79, 97.)
-

孙宇 上海图书馆系统网络中心高级工程师。上海 200031。

周纲 上海图书馆系统网络中心工程师。上海 200031。

(收稿日期:2019-09-11;修回日期:2019-10-29)