

# 清华大学图书馆一体化资源管理平台建设

## ——以ALMA系统实施为例

窦天芳 杨慧

(清华大学图书馆, 北京 100084)

**摘要:** 本文结合清华大学图书馆实际, 阐述当前图书馆资源管理与服务现状和一体化资源管理平台建设思路。选取平台建设过程中的关键问题, 从建设思路、处理原则、应用效果等角度展开分析, 希望对国内一体化资源管理平台建设和管理提供有益借鉴。

**关键词:** 一体化资源管理平台; 电子资源; 元数据; 数据迁移; ALMA

**中图分类号:** G251; C931

**DOI:** 10.3772/j.issn.1673-2286.2020.05.001

### 1 图书馆资源管理与服务现状

当前, 海量电子资源已成为资源主体, 与纸本资源、本地数字化资源交织融合, 构成全新的信息资源环境。

在新的信息资源环境下, 用户更依赖社会化搜索引擎或学术性搜索引擎查找学术资源, 并希望获得一站式、更细颗粒度、更智能的资源获取服务。图书馆高度重视电子资源的建设、揭示与推广应用。为满足持续变化的文献获取需求, 图书馆不断跟进数字内容传播机制, 建设并优化电子资源服务系统。2000年以来, 清华大学陆续建设了数据库导航系统、电子期刊导航系统、电子图书导航系统, 以及基于跨库检索的联邦检索系统 (Federated Search)、基于元数据仓储的统一资源发现系统 (Discovery Tool Based on Metadata Repository) 和链接解析系统 (Linking Resolver System)。其中, 联邦检索系统的统一检索功能因检索机制不可持续而被基于元数据仓储的统一资源发现系统所替代; 这3个导航系统经历了从静态页面到动态页面的变迁, 按资源类型指引用户获取所需。基于元数据仓储的统一资源发现系统作为电子资源搜索引擎解决了绝大多数资源的深度检索问题, 在用户界面、检索效率、检索质量, 以及对检索结果的有序处理等方面改善了用户体验,

成为图书馆高度整合所有类型资源并提供资源发现和传递服务的新系统平台<sup>[1]</sup>。

截至2019年底, 清华大学图书馆订购的电子期刊 (含收集到的OA期刊) 14万种, 电子图书2 000万册, 分布在300多个服务平台。前述服务系统引导用户检索或访问不同类型的电子资源, 年处理请求量 (Pages/Search) 超过9 000万人次, 基本满足电子资源从检索到内容获取的一站式服务需求, 初步实现了从检索到电子图书或电子图书章节、期刊、期刊论文、论文中的科学数据等多层级、细颗粒度的资源获取。

与电子资源的体量和服务力度相比, 图书馆对电子资源管理的重视普遍不足。首先, 图书馆电子资源管理方式粗放。绝大多数图书馆利用电子资源服务系统兼做电子资源的内容管理, 不涉及业务流程管理。在纸质资源管理中积累下来的管理规范尚未在电子资源管理中体现, 电子资源的内容数据和业务数据较难关联。此外, 多载体、多层级的内容数据因服务系统的不同长期处于分散管理的局面, 数据割裂, 难以呈现资源全貌。随着电子资源体量的增长, 图书馆亟需由粗放式资源管理向精细化资源管理转型, 面向灵活多样的资源服务需求提升资源管理能力。

## 2 一体化资源管理平台建设思路

随着时代进步,图书馆职能不断被赋予新内涵,但文献资源保障仍是当下以及可预见的未来图书馆最重要的服务职能。因此,探索与信息资源环境相匹配的文献资源保障路径,优化文献资源保障服务仍然是图书馆应坚持做好的基本功<sup>[2]</sup>。

### 2.1 建设思路

从资源保障方案的整体框架上看,图书馆应将面向馆员的数据管理和面向用户的服务管理从逻辑上剥离。面向馆员的是资源管理窗口,重点解决图书馆多类型资源的管理问题,整合业务数据和内容数据,汇总内容使用数据,可称之为后端系统;面向用户的是资源服务窗口,重点解决资源发现和资源获取问题,记录用户行为数据,可称之为前端系统。泛在的用户环境要求前端系统必须敏捷灵活,可根据不同场景或用户环境快速部署。此时的后端系统则要足够有序和稳健,可为不断拓展的前端服务提供高效的数据支撑。2016年,清华大学图书馆决定启动图书馆一体化资源管理平台建设,重塑面向馆员的资源管理窗口。考虑到资源管理窗口对数据依赖程度较高、业务复杂,清华大学图书馆拟选择云端平台,通过本地化研发和部署实现清华大学图书馆一体化资源管理功能。

经过选型评估,清华大学图书馆决定使用ExLibris公司的ALMA产品作为构建一体化资源管理平台的基础软件。为此,图书馆需重新梳理从资源采选到资源服务的业务逻辑,在ALMA中完成业务流程管理、数据管理,建立清华大学图书馆的业务管理规范。为保障平台建设顺利,图书馆成立了ALMA实施工作组,完成业务流程设计和系统迁移工作。

### 2.2 基础软件平台功能概述

ALMA是ExLibris公司2012年推出的云端资源管理系统,拥有庞大的元数据仓储,汇集众多资源供应商多层次的数据和系统一起部署在云端。ALMA推出早期便和北美地区多所学术图书馆合作,不断优化系统功能,目前ALMA已在图书馆服务平台市场占据主导地位<sup>[3]</sup>。2016年,ALMA正式宣布将云端平台部署在北京的数据中心(Data Center),国内图书馆开始将ALMA

列入采选范围。ALMA提供六大功能模块,包括资源订购(Acquisition)、资源揭示(Resource)、资源获取(Fulfillment)、用户管理(User Management)、统计分析(Analytics)、通用管理(Administration)。业务模块之间、模块内部各操作环节按照 workflow 管理。ALMA是基于角色(Role based)和 workflow (Workflow based)管理的软件系统,提供业务流程管理框架,图书馆可以按照系统默认的业务流程管理框架进行角色设计来塑造本馆资源管理流程;也可以自定义 workflow (Work Order),对一些独有或特定属性的工作自定义管理模式。因ALMA自带庞大的社区知识库(Community),ALMA用户可以自由选择是否共享本地资源,对清华大学图书馆而言,更重要的是可从社区知识库中获取大量的电子资源内容数据。

与传统的纸质资源管理系统不同,ALMA是一体化的资源管理系统,很多传统的概念因不再适用当前信息环境而被全新的概念所替代。如资源获取(Fulfillment)替代了传统的流通(Circulation)概念,从内涵上涵盖了传统意义上的纸质图书(Physical)借还管理、数字化(Digital)资源获取、多馆联合资源获取(Resource Sharing)和电子(Electronic)资源在线访问等事项。这种设计理念更加方便图书馆将各类(各馆藏地)资源的获取政策通盘考虑,制定的获取规则自洽。此外,ALMA将用户、馆藏地、资源、规则在逻辑上分层设计,灵活适用各种管理要求。

## 3 一体化资源管理平台建设内容

本节选择一体化资源管理平台建设过程中需重点关注的内容,从概念、流程、思路、原则等方面阐述。

### 3.1 电子资源管理

上述讨论可以发现ALMA的功能模块按照业务流程划分,而非按照资源类型划分。因此在总的管理流程上ALMA对电子资源、纸质资源及其他类型资源没有区分,都需要对订购、揭示、获取等环节进行管理,但每个功能模块下,可根据资源类型进一步细化工作流程。本节以电子资源为例说明资源管理过程。

电子资源管理和纸质资源管理相同,主要包括资源订购、揭示管理两大部分。因大多数电子资源使用量由资源商记录,因此电子资源管理中还增加了统计

数据收割的模块。电子资源订购管理包括经费、供应商、内容服务商、订单、发票等，与实体资源相比，电子资源没有实物到馆的验收环节。电子资源订单发送后，对应角色可以通过在任务清单（Go to Task List）中直接点击“完成”（Done）转到发票管理环节。在资源揭示模块，电子资源的处理流程和纸质资源不同。为阐述清楚，先对资源揭示过程中涉及的基本概念作简要说明，这些概念是在电子资源服务系统中常用的概念，也是电子资源服务生态系统中多角色对话的通用术语。电子资源库（E-Collection），通常指一个资源集合，可以理解成图书馆订购的数据库。电子资源库可以包括一个独立的数据库（Databases），如文摘索引数据库，也可以是包含全文资源的按整库订购的资源包（Aggregator Packages）或者是按刊选购的资源包（Selective Packages）。资源包中有多个单册（Portfolio），这些单册特指在该资源库里的数字对象（Object）。电子资源库和单册之间可以定义服务方式（Service），如对全文数据库而言，服务方式就是获取全文，需要配置获取全文的程序（Parser）或者从社区知识库中找到ALMA已经为该资源撰写好的全文链接程序（配置代码），这段程序完成从元数据到对象数据的链接。如Elsevier ScienceDirect是一个电子资源库，里面包含若干期刊，其中期刊Cell是一个“Object”，Elsevier ScienceDirect中的Cell就是一个Portfolio，“Portfolio”特指包含期刊收录年限范围的期刊信息。图1表示了概念间的关系，表1以示例形式对各个术语做出进一步解释。

电子资源揭示的过程是建立电子资源库、激活或导入单册数据、配置全文链接程序的过程。因ALMA自带社区知识库，馆员可以在社区知识库中选择本馆订购的资源包和与之匹配的服务方式进行全库激活或选择性激活以完成资源的揭示工作；同时ALMA支持自定义

表1 电子资源管理基本概念示例

术语	示例
电子资源库 (E-Collection)	Elsevier ScienceDirect
数字对象 (Object)	Cell
单册 (Portfolio)	1995年以来的Cell
服务方式 (Service)	全文获取
配置程序 (Parser)	CODE: CODEXXX

的数据导入规则，馆员也可以批量导入外部资源列表，快速完成资源揭示工作。

全球范围内，多数资源商根据用户访问日志生成COUNTER (Counting Online Usage Network Electronic Resources) 标准的统计数据。为进一步发挥统计数据在决策管理中的作用，图书馆须及时获取不同电子资源商生成的COUNTER数据，进行统一处理。SUSHI (Standardized Usage Statistics Harvesting Initiative) 是COUNTER格式的统计数据交换协议，可在图书馆与资源商之间搭建数据获取的桥梁，2007年成为NISO标准<sup>[5]</sup>。COUNTER从第三版开始强制要求支持COUNTER协议的资源商支持SUSHI协议，加速推动了图书馆与电子资源商之间的数据快速交换，为图书馆做好基于使用数据的管理决策提供了技术保障。在电子资源管理中，配置SUSHI协议及时获取统计数据是关注电子资源服务效果、评估电子资源平台商使用成效的一个重要手段。目前，多数资源商的统计数据可以按月收割，方便图书馆快速掌握资源利用情况。表2是以期刊Nature为统计口径，展示Nature在不同平台的使用情况。

### 3.2 业务流程与业务规范管理

一体化资源管理平台的建设过程是业务流程重塑的过程，系统角色的变化直接导致业务的重新划分，这对一个运行多年的图书馆而言是个巨大的挑战。首先，图书馆是一个操作性很强的部门，即使ALMA提供了管理框架，但落实到实际业务中，仍然需要根据不同角色细化每一个操作环节。其次，ALMA按照工作流管理，相比传统的图书管理系统更加严谨，所有操作必须是对应环节的对应角色才能完成，越权操作行为受到限制。最后，ALMA提供了全新的视角和多种工具辅助资源揭示，如通过本地化开发嵌入索书号自动生成脚本，通过配置外部数据源实现与CALIS数据的双向交

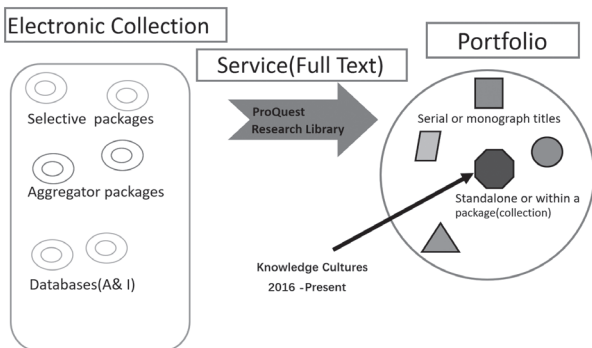


图1 电子资源管理基本概念关系<sup>[4]</sup>

表2 Nature使用统计数据

Origin ISSN	Origin Title	Platform	JR1 - Journal Usage Counter (total)	JR1 - Journal Usage Counter (html)	JR1 - Journal Usage Counter (pdf)
0028-0836	Nature	EBSCOhost	1 302	0	1 302
		ProQuest	6 069	1 892	4 177
		nature.com	726 739	146 003	72 774
0028-0836	Nature News	nature.com	0	0	0
0028-0836	Nature archive: 1869-1949	nature.com	3 435	355	774
0028-0836	Nature archive: 1950-1986	nature.com	13 005	2 096	2 400
0028-0836	Nature archive: 1987-1996	nature.com	319	40	67
0028-0836	Nature archive: 1997-2006	nature.com	80 056	11 858	8 096

换。这些系统功能的设定与人员分工密切相关，在平台建设过程中需谨慎处理。

在业务流程重塑过程中，工作组坚持以下3个原则。①传统业务能够平稳过渡。参照ALMA流程框架适当调整本馆的业务流程，尽量贴近现有部门及人员的分工。②新生业务坚持效率优先。如电子资源管理尽量按照ALMA给定的流程框架设置角色和操作规范，兼顾现有人员的分工。③自定义工作流程满足特定需求。根据本馆特定需求设置工作流程单（Work Order），进一步细化管理。目前清华大学图书馆自定义了7个工作流，解决馆藏搬迁、馆藏修复、装订等诸多事项的管理。

根据上述原则，传统业务的操作规范以部门为单位按照事项撰写，电子资源管理规范跨越部门界限按照流程编写。可以看到，这种业务规范的编写方式在图书馆实际运行中取得了较好的效果。

### 3.3 历史数据迁移

考虑到图书馆业务的延续性和资产可追溯，历史数据迁移是新平台建设无法规避的重要环节之一。需要迁移的数据应包括书目数据、馆藏数据、期刊登到数据、订单数据、名称规范数据、用户借阅历史数据等。

历史数据的迁移烦杂艰巨，难度主要表现在两大方面。一是原系统各类数据转换到ALMA的映射关系。1996年，清华大学图书馆采用INNOPAC/Millennium系统作为馆藏文献资源管理系统，提供馆藏目录检索系统。INNOPAC/Millennium系统积累了大量历史数据，包括实体资源书目记录308万种、单册记录413万册、期刊登到记录4 000条、订单记录62.18万条、名称

规范记录5.8万条等。迁移前需要根据ALMA数据结构和业务逻辑制定数据转换规则，部分既有数据需要做预处理进行拆分、合并或格式转换；同时，需要制定INNOPAC/Millennium数据提取方案，包括数据提取范围、事项及顺序等。为保证数据迁移的准确性，在数据正式迁移前要经过多次预装载测试，反复调整数据预处理方案和新旧系统数据映射规则。二是对历史数据的清洗。人工著录的数据会因编目规则的变化或人为因素导致格式不完全一致，需要对INNOPAC/Millennium系统数据进行全方位地梳理、排查并逐一修正相关问题，保障既有数据的格式统一。数据清洗及校验工作需要经验极为丰富的馆员主导。另外，原系统还如实记录了历史留存的各类问题，如系统挂起的丢书、欠款、书目禁忌等问题，需要在数据迁移前给出处理方案。

为最大限度地减少因数据迁移给用户带来的不便，数据迁移分两步进行。第一步是冻结馆藏数据进行迁移，馆藏迁移完成后由图书馆对数据进行校验，如校验结果符合预期则启动第二步，即冻结用户数据进行用户数据的迁移。经过缜密设计，清华大学图书馆基本上将所有类型的历史数据成功迁移，读者的借阅数据在新旧两个系统间平稳过渡。

### 3.4 用户数据管理

用户管理是资源管理系统的重要模块。清华大学图书馆包括一个主馆和若干专业分馆及院系资料室，总计13个相对独立的实体，虽然清华大学图书馆致力于持续推动全校通借通还制度，但受资源建设经费和资产管理的限制，主分馆针对不同人群制定了不同的借阅规则。ALMA系统需根据现行的借阅规则细分人群

并兼顾统计分析的需要,最后工作组讨论确定了50个用户组(User Group),并完成授权管理和流通规则设置(Fulfillment of Users)。

用户的基础数据要从学校多个用户库同步到ALMA系统,按照身份自动匹配到用户组。用户认证则通过调用学校提供的统一身份认证系统来实现。

## 4 一体化资源管理平台运行优势

2017年10月5日,一体化资源管理平台正式上线,同时图书馆所有的采编、流通业务全部迁移到新平台进行,与新平台同步升级的资源发现系统“水木搜索”同时升级上线。系统运行以来,充分展现了一体化资源管理平台的运行优势,主要表现在以下两个方面。

### 4.1 业务数据和内容数据深度关联

ALMA作为一体化资源管理平台,可针对全媒体资源进行全业务流程的管理。从宏观层面讲,系统可以将资源经费、资源类型、资源用量关联起来,形成实时报表的同时可以给出带有智能推荐的决策意见。针对某一特定类型的资源,如电子资源可按照题名、资源包、代理商、平台、经费等多个维度生成报表,甚至可按照成本高低给出资源清单,帮助图书馆监测经费使用绩效和细颗粒度的资源的利用状况。因涉及经费等敏感数据,笔者选择部分期刊展示篇均成本(见表3)。当然,系统自带图形化报表界面(Dashboard),方便馆员直观了解数据表现(见图2)。需要说明的是,成本最低不代表用量最高,需有管理者综合使用各类数据做出业务监测和优化。这些细化到期刊的效益分析没有强大的管理系统是难以想象的。

表3 部分期刊篇均成本

Title (Normalized)	Electronic Collection Public Name	Platform	Cost Per Use
Physical Review Letters	American Physical Society Journals	APS	0.03
Physical Review B	American Physical Society Journals	APS	0.04
Physical Review A	American Physical Society Journals	APS	0.09
Cell	Elsevier SD Cell Press	ScienceDirect licensed content	0.11
Journal of The American Chemical Society	American Chemical Society Journals	ACS Publications	0.13

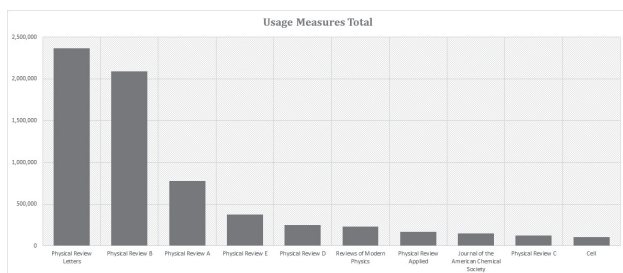


图2 期刊用量情况

### 4.2 内容数据统一管理

内容数据与前端应用在逻辑上分离是大势所趋,一方面有利于内容数据的统一管理与维护,另一方面也极大地方便了前端展示系统的灵活部署。

一体化资源管理平台已经对包括电子期刊、数据库、电子图书、纸质图书在内的各类内容数据进行集中

统一管理,方便馆员随时掌握资源全貌和同一资源不同载体的分布及重合情况,为精细化的资源管理提供了强大的平台保障。在资源管理的操作方面,ALMA提供了多种便捷的功能,如动态和静态结果集、可灵活配置的外部数据源、在数据导入或更新时可方便使用的标准化规则、名称规范数据与书目数据的关联等,有效地保证了数据处理的效率和质量。此外,ALMA具有很好的开放性,提供多种API供第三方应用调用,所有用于资源导航的前端系统都可将ALMA视作数据平台。2017年10月以来,清华大学图书馆的电子图书导航、电子期刊导航系统相继升级,将底层数据从各自为政的系统平台转移到全部调用ALMA平台数据,有力保障了资源揭示的效率和数据一致性。未来,ALMA作为单一数据平台还可进一步拓展移动端的服务渠道,为前端的敏捷应用提供稳健的数据支撑。

## 5 一体化资源管理平台运行挑战

截至2020年1月,全球有1 900余家图书馆采用ALMA作为一体化资源管理平台。ALMA立足于全球资源管理与传播环境研发而成,又充分吸收了北美地区图书馆的管理经验,在角色管理、业务流程管理等方面理念先进,逻辑严谨。系统默认的业务流管理框架给我国用户提供了非常好的业务管理思路,但是如何将系统的管理优势转化成业务管理优势还有很长的路要走,需要图书馆在文化、理念、机制上做好相应准备,保持开放务实的心态,逐步建立科学规范的管理机制,扎实做好资源管理工作。

ALMA提供了丰富的统计功能和图形化报表展示,涵盖常见的图书馆业务统计和活动监测。但是ALMA的统计分析模块基于Oracle Business Intelligence构建,在实时决策和活动监测方面的应用对馆员能力要求较高,图书馆需在Oracle Business Intelligence方面加大投入,为数据敏感的管理决策做好人力准备。

ALMA部署在云端,按月升级并发布功能变更报告。功能变更报告包括已修复的系统问题清单和新增功能清单,方便馆员了解系统最新状况。其中,2019年4月推出的“View Recommendations”功能给图书馆带来智能管理,可针对馆员操作给出推荐意见,如馆员手工上传COUNTER数据时,系统会在View Recommendations中提示可通过配置SUSHI完成统计数据的收割而无须手工上传。这种云端系统对保持技术的快速迭代优势明显,图书馆应适时调整思路,更专注于内容建设和适用性技术研究,在专业细分的服务生态环境中突出重围、脱颖而出。

## 6 总结与展望

自2017年10月5日清华大学图书馆一体化资源管理平台上线至今,清华大学图书馆的资源管理流程不断优化,操作模式基本稳定;电子资源订购逐步规范,电子资源揭示效率大幅提升;系统与业界主要的联机编目系统无缝关联,纸质资源揭示效率稳步提升。此外,一体化资源管理平台支持多种流通政策,系统运行以来也拓展了多种服务窗口,与新生入学开通借书权限、毕业生离校、证卡管理等系统对接,在馆内外收到较好的反馈。

ALMA是在元数据开放关联的基础上以业务流管理的思路开发而成的资源管理系统,是支持与外部环境进行广泛互联的生态系统。ALMA的规模化应用正在改变知识传播及知识服务格局,其在我国的落地应用必将对我国图书馆行业的软件研发、数据建设、服务理念带来难以估量的深远影响。

### 参考文献

- [1] 窦天芳,姜爱蓉. 资源发现系统功能分析及应用前景 [J]. 图书情报工作, 2012, 56 (7) : 38-43.
- [2] 陈建龙. 大学图书馆现代化转型发展刍议 [J]. 大学图书馆学报, 2020, 38 (1) : 5-12.
- [3] 曾建勋. FOLIO的启示 [J]. 数字图书馆论坛, 2019 (10) : 1.
- [4] ALMA Online Help [EB/OL]. [2019-11-26]. <https://knowledge.exlibrisgroup.com/Alma>.
- [5] 陈大庆,叶兰,丁培. 电子资源使用统计收割标准SUSHI的实施与应用研究 [J]. 中国图书馆学报, 2018, 44 (2) : 46-60.

### 作者简介

窦天芳,女,1973年生,硕士,副研究馆员,研究方向:电子资源管理、信息系统, E-mail: ddoutf@tsinghua.edu.cn。  
杨慧,女,1964年生,本科,副研究馆员,研究方向:文献编目。

Unified Resource Management Platform Construction of Tsinghua University Library: Taking the Implementation of ALAM as an Example

DOU TianFang YANG Hui  
( Tsinghua University Library, Beijing 100084, China )

Abstract: According to the current status of library resource management and service, a solution of Unified Resource Management Platform is introduced in this paper. Taking ALAM implementation in Tsinghua Library as an example, some key issues, such as data migration and e-resources management are summarized.

Keywords: Unified Resource Management Platform; Electronic Resource Management; Metadata; Data Migration; ALMA

(收稿日期: 2020-04-12)