

康奈尔图书馆 FOLIO 案例研究*

詹·柯尔特 黛布拉·豪厄尔 (康奈尔大学)

嵇婷¹ 张春景¹ 胡小菁² (译) (1 上海图书馆 2 华东师范大学)

摘要 康奈尔大学是一所私立大学, 由纽约州政府拨地建立, 拥有公共使命。在使用了 20 年 Ex Libris 的 Voyager 作为图书馆集成系统 (ILS) 后, 康奈尔大学图书馆于 2021 年 7 月 1 日完成了向开源平台 FOLIO 的迁移。本案例研究记录了康奈尔图书馆从选择 FOLIO 到上线的历程。

关键词 康奈尔图书馆 案例研究 FOLIO

Cornell Library FOLIO Case Study

Jenn Colt, Debra Howell (Cornell University)

Ji Ting¹, Zhang Chunjing¹, Hu XiaoJing² Trans. (1 Shanghai Library; 2 East China Normal University)

Abstract Cornell University is a private university with a public mission as the Land Grant institution for New York State. After 20 years using the vended application Voyager by Ex Libris as the Library's integrated library system (ILS), on July 1, 2021 Cornell University Library completed their migration to the open-source platform, FOLIO. This case study chronicles the library's journey from FOLIO's selection to go-live.

Keywords Cornell Library, Case Study, FOLIO

1 通往 FOLIO 之路

康奈尔大学是一所拥有公共使命的私立大学。康奈尔大学的学生人数约为 25,000 人, 校园用地由纽约州政府提供, 是一所私立捐赠大学, 是常青藤盟校 / 古八校的成员, 也是纽约州立大学的合作伙伴。康奈尔图书馆秉承康奈尔大学的使命, 拥有 20 个不同的实体和数字图书馆, 850 万册藏书和 170 万本电子书, 约有 400 名员工。

在使用 Ex Libris 公司的图书馆集成系统 (ILS) Voyager 20 年后, 2021 年 7 月 1 日, 康奈尔图书馆迁移到新的开源平台 FOLIO——The Future Of Libraries Is Open, 意为图书馆的未来是开放的。

康奈尔大学在 2010 年代早期就开始探索

替换 Voyager 系统的选择。由当时的负责信息技术的图书馆副馆长和首席技术战略家 Dean Krafft 领导的一个委员会, 与哥伦比亚大学的同事一起分析了可选的 ILS 环境, 该项目被称为 2CUL 项目。2CUL 还包括其他一些前瞻性的想法, 如共享馆藏开发和技术服务整合 (Harcourt & Leblanc, 2017)。学校想要一个能够促进深度合作并能引领新方向的系统。在评估了 Alma 之后, 学校决定避免过度依赖单一的供应商, 因此便开始关注 Kuali 开放图书馆环境 (OLE)。学校认为 OLE 最契合对开源的承诺, 以及馆员直接参与开发的希望。在决定加入 Kuali OLE 后不久, 社区做出了一个重要的决定, 从根本上改变项目的方向, 并停止了 Kuali 平台, 重建代码库。这个变更诞生了

* 本文根据《国际图书馆学报》(International Journal of Librarianship, 简称 IJol) 2021 年 12 月 FOLIO 特刊发表的英文论文做全文翻译。原文出处: Colt, M., & Howell, D. (2021). Cornell Library FOLIO Case Study. International Journal of Librarianship, 6(2):13-20. <https://doi.org/10.23974/ijol.2021.vol6.2.205>.

FOLIO，它是“The Future of libraries Is Open”的缩写。因此，康奈尔图书馆在 2016 年 FOLIO 社区成立之初开始参与。

编目与元数据服务主任 Jason Kovari (2021.8.30 个人交流)指出：“我们决定实施 OLE 是因为对开源有信心。我们厌倦了与供应商绑定在一起，更厌倦了在系统开发中缺乏代理，这对实现图书馆的功能至关重要。而且，这是支持一个不断发展的系统的一种手段，该系统跟随需求一起发展。FOLIO 就是建立在这种理念之上的。我们不是因为任何一组具体的功能而去实施 FOLIO；相反地，我们选择 FOLIO 的原因是我们能决定开发的方向并拥有一个非常优秀的同行社区。”

FOLIO 提出后不久，康奈尔大学就开始贡献人力、物力和财力来帮助 FOLIO 的建立。学校坚持发展和实施 FOLIO 的最大因素包括：开放获取的愿景、微服务架构、图书馆核心管理功能、批量化生产流程、能够访问和修改源代码、多种元数据选项（MARC、非 MARC、关联数据等），以及电子资源管理（ERM）功能的集成。

2 实施

经过几年的 FOLIO 开发，2019 年康奈尔大学成立了一个团队来管理 FOLIO 的实施，目标是在 2020 年 7 月上线。学校组建了一个实施团队，成员来自图书馆各个关键部门，包括：报告、财务、元数据管理、访问服务、用户测试、采访、连续出版物处理、编目、电子资源管理（ERM）、培训、基础设施、集成、发现和数据迁移。除了这些领域的“引领”之外，许多图书馆工作人员成为专题事务专家（SME），帮助培训，或参加 FOLIO 社区特别兴趣小组（SIGs）。图书馆还聘请了一名信息技术项目经理来领导这个项目。

康奈尔大学需要做的第一个决定是决定是否与供应商签订托管合同还是自己管理 FOLIO。在适当的资源条件下，图书馆可以自己管理、维护和升级 FOLIO。然而，很明显，图书馆没有足够的人员来自行管理。经过多方分析，康奈尔大学选择了由 EBSCO 提供托管

和迁移服务。EBSCO 的实施顾问是迁移和实施过程中宝贵的合作伙伴。FOLIO 最大的好处之一是，它可以通过设置完全定制，而 FOLIO 最困难的事情之一也正是它可以通过设置完全定制。实施顾问在整个过程中提供了支持、指导和解决方案，帮助项目选择最适合的功能、设置和选项。

实施一个仍处于开发中的系统需要应对一系列特有的挑战。系统实施是否成功依赖于 FOLIO 的发展是否保持正轨。2019 年 12 月，团队意识到，需要评估 2020 年 7 月上线 FOLIO 的可能性。FOLIO 开发使用 Jira 来跟踪所有开放的功能特性（features），因此团队创建了一个 Jira 仪表盘，它对管理实施至关重要。Jira 仪表盘显示了在上线时、上线后一季度、上线后两季度时需要的功能排名。为了评估在 2020 年 7 月上线的决定，团队遍历了上线时需要的所有功能排名，并决定它们是否是必需功能，还是短期内非必要功能，或是可用其他方案代替的功能。团队一个指导原则是图书馆工作人员方面的一些问题可商议处理，但所有面向用户的功能必须正常工作。2019 年 12 月，团队最终发现有 22 项必备功能在 2020 年 7 月前无法完成，因此决定推迟一年至 2021 年 7 月上线。

与此同时，康奈尔还将 ERM 系统从 ProQuest Intota 迁移到了 FOLIO ERM。这个项目由康奈尔的 Peter McCracken（采访和电子资源战略图书馆员）和 Emma Raub（专题事务专家、电子资源图书馆员）主持。该项目能够继续推进的原因是它是完全独立的，而当条件成熟时，它又可以被完全整合到 FOLIO 中。2020 年 1 月，康奈尔成功实施了 FOLIO ERM。

在决定将上线时间推迟到 2021 年之后，团队集中精力解决这些必备功能项。许多必备功能都在 FOLIO 开发计划中。团队密切跟踪这些功能，如果这些功能有被变更优先顺序的可能性，团队就会指出它们的重要性。对于其他的必需功能项，团队会拓展思路找一些其他的方法。团队与芝加哥大学、德克萨斯农工大学和杜克大学合作，开发了在线计算机图书馆中心（OCLC）- 单记录 - 导入功能。团队承担

了核算系统自动输入和流通日志自动导出功能,并将代码贡献给了 FOLIO 社区。

康奈尔图书馆已经定制集成了许多其他应用程序,见表 1:

表 1 康奈尔图书馆集成的应用程序

功能	产品
大学财务系统	Kuali(开源)
康奈尔核算系统	
资源发现与获取	Blacklight(开源)
馆际互借	BorrowDirect 和 ILLiad
图书馆技术服务	全球在线书目信息 (GOBI)
图书馆技术服务	EBSCO 发现服务 (EDS)
图书馆远程存储	CaiaSoft
教学参考书	Ares
珍本和手稿收藏	Aeon
珍本和手稿收藏	ArchiveSpace

图书馆技术服务 (LTS) 工作流程 内部定制开发的应用程序

其中许多集成功能需要开发。有些功能由 FOLIO 项目开发,有些由供应商开发,还有一些由内部开发。团队分配了大量的时间来测试系统的集成。然而,计划上线的 FOLIO 版本推迟发布了两个月,团队仅剩大约三周的时间来完成集成和测试。因此,当进行系统切换时,一些集成组件不能工作。例如, BorrowDirect 适用于借书,但不适用于还书。团队在上线后的几周内完成了集成工作。

2020 年 3 月,全球新冠疫情让大多数图书馆员工在接下来的一年里只得待在家中。FOLIO 的剩余实施工作和上线切换工作只好远程进行,或根据情况在混合环境中进行。因为 FOLIO 是一个全球社区,在大多情况下,人们已经习惯了远程工作。其他方面,例如在员工培训与参与方面,也受到了很大的影响。

考虑到大部分的培训是线上进行的,从 2021 年 2 月开始,学校开始进行一种分层的培训方法,为上线做准备。作为每个发布周期的一部分, FOLIO 开发社区会举办一次捉虫大会 (Bug Fest)。在捉虫大会期间,大家对发布的所有特性功能进行测试。康奈尔将捉虫大会作为一种让员工参与 FOLIO 的方式。许多早期的捉虫大会测试人员已经成为 FOLIO 专题事务专家,他们帮助图书馆中的其他人解决问题。学校创建了一个培训团队,负责

进行 FOLIO 演示和举办特定功能的培训课程,由于疫情,所有这些都必须通过 Zoom 完成。团队对每个环节都进行了记录并在网上提供。同时还创建了包含常见问题和趣闻趣事的 Confluence 网页,并每周通过电子邮件发送项目新闻和培训信息。最后,学校决定在上线前 10 天停止大部分流通功能,让资源服务工作人员有机会休息,并不断进行 FOLIO 培训。

在康奈尔大学上线 FOLIO 的两年前,数据迁移工作就开始了。学校采取了一种团队导向的方法,即数据迁移作为康奈尔实施团队的一部分。有两名程序员也在数据迁移团队工作,一名专注于书目数据,另一名专注于采访、用户和流通数据。因为康奈尔与 EBSCO 合作,有了托管商,学校能够专注于从 Voyager 提取数据映射到 FOLIO,而 EBSCO 则负责将数据加载到 FOLIO,因为他们能更直接有效地访问数据库。

下面列出了项目在数据迁移工作中的一些经验与教训。

(1) 从清理数据开始

在迁移过程初期,图书馆技术服务部门的一个小组负责启动数据清理工作。该小组专注于清理书目、馆藏和单件记录。由于 FOLIO 在典藏中有一个分面导航和检索系统,因此,清理书目和馆藏数据以增强检索能力、改善 FOLIO 典藏展现能力是很重要的。

书目信息是最大的待清理数据池,其他领域也有相应工作。财务部门精简了图书馆的资金结构,公共服务部门简化了流通规则,采访部门清理了不必要的采购订单。这些清理工作带来了额外的收益和机遇。例如,财务清理将单个经费合并到团队经费中,使得康奈尔大学图书馆在转移分配方面具有了更大的灵活性。

(2) 与数据消费者密切合作

康奈尔大学的数据迁移团队定期与图书馆的各个部门会面,了解他们的工作。虽然用户体验可能看起来不像是数据迁移的一部分,但深入研究数据使用人员的工作流和行为,可以做出如何在 FOLIO 框架下迁移数据的明智决定。

理解基于数据支撑的工作流有助于确定以

何种形式迁移哪些数据。通过与用户广泛接触,鼓励他们探索 FOLIO,团队帮助他们了解在新系统中的状态,这样他们就可以提供基于数据需求的更多更好的信息。

除了揭示哪些数据需要迁移以及迁移到哪里之外,团队还与用户讨论了哪些数据不需要迁移。例如,在采访中,无需迁移已关闭的采购订单,这样既减少了需要迁移的数据量,又减少了所需的清理量。在流通数据方面,仅需迁移未结借阅以及涉及罚款和费用的借阅数据。

因为 FOLIO 的流通规则较 Voyager 系统规则大大简化,迁移团队还研究了在迁移期间如何更改构成流通规则(如借阅政策和单件类型)的数据碎片,以使 FOLIO 流通规则起作用。

在这项工作中,参与 FOLIO 社区是非常有必要的。团队参加了数据迁移特别兴趣小组,并组织了几次与其他从 Voyager 迁移到 FOLIO 的图书馆的对话。FOLIO 社区是数据迁移成功的关键支撑,这与它在项目中的其他领域所发挥的作用一样。

(3) 记载数据映射文档

数据迁移自然需要大量的文档。团队创建了从 Voyager 迁移到 FOLIO 的数据映射电子表格。

FOLIO 中保存了每个实体的电子表格,并将数据元素从 Voyager 字段映射到 FOLIO 字段。团队记录了如何从 Voyager 中提取每个数据点。这是必要的,因为虽然 FOLIO 中的某些字段是某个实体的一部分,但它们是从 Voyager 中的多个实体中提取的。还有一些字段在 Voyager 中并不存在,因此它们的内容必须基于 Voyager 中的各种数据点来推断,以便映射到 FOLIO 中。

(4) 制定多迭代迁移计划

团队在上线前几个月就开始进行 FOLIO 数据迁移测试。这样做可以发现遗漏的数据问题。将原有数据迁移到 FOLIO 也是帮助员工理解使用 FOLIO 感受的一个关键环节,因为他们能看到数据中的漏洞或问题。

托管商为员工配置了可访问的沙盒系统后

便将数据加载到沙盒中,然后员工就可进行测试和检查,并给予反馈。虽然报告和日志可以揭示数据中的模式和问题,但实际用户对数据的手工检查在发现忽略问题和寻找初始映射中的数据点方面是极其有用的。

(5) 提供迁移报告和日志记录

EBSCO 提供了每次迁移迭代的报告,这样在出现问题时可进行修复。日志记录需要包含允许用户通过提取、转换和加载过程跟踪数据的信息。确保在 FOLIO 的所有数据中嵌入来自 Voyager 的标识符,这样既可以维护数据之间的关系,也可以跟踪可能发生的问题。

每次迁移迭代都会产生大量需要分析和评估的数据。有时日志所显示出的 bug 需要向 FOLIO 社区报告并进行处理。有时日志记录会指出 Voyager 数据中存在问题,需要加以纠正。还有的时候,FOLIO 所需的数据点 Voyager 没有提供,需要特别为迁移生成。在某些区域,当 Voyager 没有提供所需的数据点时,需要设置字段的默认值。

(6) 周密计划切换日程

通过迭代迁移,团队能计算出数据加载到 FOLIO 所需的时间。这有助于计划停用 Voyager 和开启 FOLIO 之间的时间间隔。团队创建了一个操作顺序,它记录了在转换期间各种类型的数据何时加载,并确保在开放新系统给员工和公众使用之前,将这些数据准备到位。一些数据迁移工作,尤其是那些不属于流通或发现的数据迁移工作甚至在线上后仍在继续。

在 FOLIO 上线后,继续迁移和清理数据用了大约 6 周时间。当时数据迁移已成功完成,但 FOLIO 中的数据仍需继续细化。方法包括认真倾听 FOLIO 用户的需求,仔细考虑 Voyager 和 FOLIO 数据模型之间的差异,以及与托管商进行合作。

3 项目管理的挑战与建议

实施开源系统,尤其是一个仍在发展和进化中的系统,沟通是最重要的项目管理工具;有必要在无条件的支撑项目和不要带来太多负面作用之间取得平衡。康奈尔大学实施项目经理

每月举行一次“FOLIO 周五”会议, 议程包括项目进展、时间表和提问环节。我们发现, 工作人员更加欢迎对项目健康状况的诚实评估, 以及交流那些受影响领域的事实细节。

此外, 有效的风险管理必不可少。你可能会觉得自己在不断地灭火。然而, 这是管理实施仍在开发中的系统所不可避免的。使用功能需求列表来进行差距分析是不可能的。FOLIO 实施是一个大型的、复杂的项目, 存在很多变化。根据经验, 建议培育一种去中心化但统一的方法, 同时为每个功能领域确定一个引领者, 项目负责人必须相信每个领域的引领者都为康奈尔大学和社区做出了最好的决定, 并且会报告每个重大问题。

(1) 上线

2021 年 7 月 1 日上线前的几周是一片混乱。几个必备功能处于无法上线发布的边缘, 员工们对超过一年的疫情感到疲劳, 对 ILS 的切换排在最高优先级有一定的情绪, 而和康奈尔一样原计划同时上线的大型图书馆都推迟了上线。实施小组不断对项目进行重新评估, 图书馆执行小组评估项目影响, 并给予了持续的指导和领导。

一个接一个地, 团队解决了一项项关键问题, 完成了集成和数据迁移, 并且通过 Zoom 会议、电子邮件和康奈尔图书馆 Slack 频道与图书馆工作人员保持着持续的沟通。

图书馆完成向 FOLIO 的变迁不仅是康奈尔大学的一个里程碑, 也是全球图书馆社区的一个里程碑。康奈尔大学是世界上第一个使用 FOLIO 的大型研究图书馆, 其他图书馆正在从康奈尔的经验中学习。在实施如此规模和复杂的软件系统时, 能够获得非常流畅的上线体验, 这要归功于图书馆各个部门工作人员的奉献。

参考文献

- [1] Kate Harcourt, Jim LeBlanc. Finale and Future: The 2CUL Technical Services Strategic Alliance[J]. Library Resources & Technical Services, 2017(1):43-55.

(2) 上线之后

上线后, 学校专注于保持员工士气, 并确保所有面向用户的功能正常运行。大家理解、接受系统切换使得一些员工的职能将会受到影响, 并对此进行交流。例如, 在上线后, 技术服务人员报告说, 他们的工作效率约为 20%。几周后, 这个比例上升到 40%。预计今年剩下的两个 FOLIO 版本将会在功能和生产力上有一个剧增。

为了继续保持透明度, 团队在“已知问题和状态” Confluence 页面上跟踪所有的 bug 和功能缺陷, 所有工作人员都能访问到。实施团队正在减少, 每周的会议仍在举行, 以便让工作人员询问有关 FOLIO 的问题。康奈尔图书馆 Slack 频道已经成为图书馆工作人员发布问题和接受同事帮助的一个充满活力的地方。

上线后, 大家对实施过程进行反思, 归纳了如下要点: 透明沟通的必要性、灵活与迭代的重要性, 以及让所有员工参与到过程中的价值感。

截至 2021 年 11 月, 康奈尔大学仍在运行 FOLIO Iris 版, 这是 7 月上线的版本。如果测试显示没有关键性功能 bug, 系统将升级到 FOLIO 未来发布后的各版本。

(3) 开源图书馆平台的可持续性

负责信息技术的图书馆副馆长 Simeon Warner 说, 社区支持和合作对图书馆来说比以往任何时候都重要, 特别是在共享日益增长的电子资源方面。“像康奈尔这样的图书馆从来都不是单独运作的, 认识这点很重要,” 他说, “现在, FOLIO 为我们提供了一个基础, 以图书馆合作的方式向前迈进——这令人兴奋。”

随着与 FOLIO 一起成长和发展, 我们可能永远不会迁移到另一个 ILS。

詹·柯尔特 康奈尔大学图书馆技术服务自动化与元数据系统负责人, 在图书馆 IT 和元数据技术等多个领域工作超过 15 年。长期以来一直致力于提高图书馆知识发现及开源软件社区适应性。