

美国普渡大学进修报告

王辉

中国科学院文献情报中心 学科咨询部

2014. 7. 10

进修机构：美国普渡大学图书馆 (<https://www.lib.purdue.edu/>)

指导老师：Michael Witt (<http://oldsite.lib.purdue.edu/research/witt/>)

进修时间：2014. 1. 13-2014. 7. 7

进修内容：科学数据管理服务

(一) 普渡大学的图书馆简介

普渡大学的图书馆系统包括多个学科的档案和特色馆藏：健康与生命科学；人文社会科学，教育，商业；物理科学，工程和技术。此外，还包括馆藏管理，流通和存储，资源共享，资源服务，指导和数字项目服务，分布式数据管理中心。普渡大学出版社和大学版权办公室的工作也集成在图书馆的业务中。

普渡大学的图书馆的优势与大学的核心优势一致，主要在农业、科学与工程、商业、医药、兽医。印本馆藏历史悠久、宽度和广度都支持核心学科。图书馆收藏普渡大学早期的太空探索飞行的档案和特殊馆藏文件，包括 Neil Armstrong 和 Gene Cernan 文件。Goss 收集的材料反映了从第十六到二十世纪工程和运输历史。精神药物的研究馆藏支持脑受体吸收药物药理研究控制帕金森氏症和阿尔茨海默氏病的研究。

普渡大学致力于为学生提供回答问题和形成解决方案所需的正确信息。本着这一精神，普渡大学图书馆展开全面的信息素养教育，包括在 2008 建立第一个信息素养讲座，信息素养通过图书馆学科和院系之间的合作开展。普渡大学是致力于信息素养的新兴方面，包括数据读写能力，这对研究人员管理科学研究产生的海量数据至关重要。

普渡大学图书馆学科与学科教师合作开发数据发现、访问和共享的方法和知识库。普渡大学的图书馆数据监护通过研究人员和图书馆之间的对话了解数据管理的需求。普渡大学图书馆在 datacite 和分配数据集的数字对象标识符 (DOI) 方面在国际上起到领导作用。

图书馆的空间概念发生了改变，自 2009 年以来，管理和经济学图书馆发生了一个重大改革，建立教室的自然延伸的空间，学生可以使用智能板，双向通信

设备, 图像采集设备和计算机 POD 工作站与团队协作或个人单独工作。Hicks 图书馆正在研究阶段, 并规划与科学与工程图书馆合并。

(二) 普渡大学数据管理服务状况

研究和数据服务提供整个数据生命周期对数据描述、工作流程、管理、发现、传播和保存的支持和指导。把生命周期分成三段介绍数据服务。

准备数据计划: 从事新研究前, 学科专家结合研究和数据服务, 提供数据管理计划咨询和参考资料, 完成资助者要求数据管理计划。帮助确定数据发现、传播和保存。帮助确定元数据标准的描述和访问, 促进数据共享、使用和重用。

项目过程中: 通过咨询中心、实验室、和个人的教员和学生, 提供数据服务, 满足当前正在进行的研究项目的特定需要。普渡大学图书馆的数据管理可以帮助识别当前数据工作流的需要, 通过管理解决“上游”更好的准备数据和“下游”数据分布、传播和共享问题。

项目后期: 研究项目后期需要提供数据, 可以帮助识别国家、区域的库来存储数据。学科专家结合研究和数据服务, 为用户提供有关格式、禁忌、权利和描述的咨询, 方便其他用户使用、重用或派生出新的数据。

1) 开发和发布数据

普渡大学的研究存储库提供了一个在线的、协作的工作空间和数据共享平台, 以此支持普渡大学的研究人员和他们的合作者的数据管理需求。

2) 数据管理咨询服务

在研究期间和项目后帮助用户识别新的或替代方法数据管理、工作流和传播。

3) 数据研究咨询与协作

应用图书馆学解决生物燃料、干旱、教育、农业、水质等其他学科的相关问题挑战, 可以通过 research@lib.purdue.edu 请求协作。

4) 数据引用

利用数字对象标识符 (DOI) 发布数据集, 使人们可以更方便地引用数据, 增加研究工作贡献度, 普渡是数据注册的国际机构 DataCite 创始成员之一, 可为美国机构提供 DOIs。为数据集分配的 DOIs 发布在 PURR 上。

5) 查找数据管理资源

图书馆编制各种学科领域的的数据管理研究帮助、主题指南和有价值的资源。

6) 图书馆学专业

普渡大学图书馆教师从事各种领域的研究, 包括数字图书馆、信息素养、图书馆员新角色、研究数据的生命周期、学术交流、教学。普渡大学图书馆教师需要经常提供图书馆主题的咨询、呈现和合作。

7) 找到数据存储库

Databib 是一个可搜索的目录、注册、参考书目的研究数据存储库。可以用它来找到一个存放数据的合适的地方, 以及找到其他研究数据存储库。去 DatabibFind data repositories.

8) 创建数据管理计划

使用 DMPTool 创造现成的数据管理计划。这个 DMPTool 提供数据管理计划的指导, 并协助开发数据管理计划, 指导具体的资助机构。普渡大学图书馆教师可以提供数据管理规划咨询适合您的提交。普渡大学是 DMPTool 指导委员会的特许成员。

9) 分布研究、分布式数据、分布式的方法、不同的解决方案

在多个领域(从项目到数字人文), “海量数据”已经为研究者带来了问题, NEH、NSF 等基金认为图书馆科学对帮助科研人员解决问题起到至关重要的作用。分布式数据管理中心调查研究数据的相关问题, 并与研究人员合作开发解决方案, 包括: 在分布式环境中的研究数据管理、传播和保存。

10) GIS 服务

专业服务包括: 应用程序开发, Geodatabase 设计、数据采集、数据转换/开发、地理编码、地理信息处理、项目协作、客座演讲, 指导和培训、空间分析。

(三) 在普渡大学期间工作学习情况

1) 中科院文献情报中心学科化服务与新员工培训报告

在普渡大学交流访问期间, 应普渡大学图书馆馆长 James L. Mullins 博士的要求, 面向全馆进行了关于“中科院文献情报中心学科化服务与新员工培训”的报告, 分别与普渡大学新员工培训计划工作组、学科馆员、科学数据服务团队成员、科学数据素质教育工作组、Hubzero 工作组、PURR 团队成员、就他们的日常工作、岗位所需知识与技能、服务现状、未来机遇与挑战进行交流, 完成普渡大学交流学习报告(英文)。

2) 科学数据素质教育课程学习

参与普渡大学面向研究生的科学数据素质教育课程学习, 了解数据素质教育课程框架及内容, 学习与数据管理相关的知识与工具。

3) 小科学数据共享与机构基础设施研究

通过调研国外“小科学及分支学科在数据管理实践、共享需求”项目研究小科学数据特征及数据管理实践的学科差异、共享需求, 调研国外国家层面、组织

层、大学层面的科学数据服务案例，遴选六个具有代表性的大学机构基础设施，分析这些基础设施适应广泛的小科学数据特点和共享的做法，分析机构利用基础设施服务发展面临的问题与挑战，总结可供借鉴的实践经验，完成《小科学数据共享与机构基础设施研究》。

4) 普渡大学研究仓储服务案例研究

通过嵌入 PURR(<https://purrr.purdue.edu/>)团队，参与 PURR 的团队例会、用户交流、数据上传、平台利用，对普渡大学研究知识库案例进行全面剖析。分别从 PURR 平台的建设背景，PURR 的开发平台，建设目标，数字保存政策，元数据解决方案，数据集保存战略计划和保存策略， workflow，工作机制以及 PURR 支持的大学科学数据管理服务等方面进行分析。完成《普渡大学研究仓储服务案例研究》。

5) 参与 Databib 项目

对中国数据仓储进行扫描，描述中国数据仓储的基本信息，并将其加入到 Databib 中，完善中国范围内的数据。

6) 参加 ORCID 会议

参加在芝加哥举行 ORCID 的年度会议，学习来自数据库商、大学图书馆、业协会等不同角色对 ORCID 的应用案例，与“ORCID 与机构知识库的结合案例”的报告人进行交流。会议 ppt 和 poster 链接如下：

<https://orcid.org/content/orcid-outreach-meeting-and-codifest-may-201>

(四) 收获与感受

普渡大学在科学数据管理方面和信息素质教育方面值得我馆学习。普渡大学已将数据读写能力加入到信息素质教育内容中。

普渡大学图书馆学科与学科教师合作开发数据发现、访问和共享的方法和 PURR。PURR 参考多个标准进行建设，支持数据服务，普渡大学作为数据注册的国际机构 DataCite 创始成员之一，支持每个在 PURR 上发布的数据集都可分配的 DOI，有利于数据的引用。PURR 作为先驱性的数据管理工具，其在开发及推广过程中积累的经验为我国开展数据管理实践提供了可供借鉴的参考。

在团队管理方面，PURR 团队充分利用 PURR 提供的项目共享空间，一方面直接获取用户体验，另一方面对科学数据管理相关的文件、团队会议记录做了良好

的备份与管理，供长期使用。