

# 伦敦大学城市学院图书馆数据素养教育课程剖析

郑新 李俊霞

(聊城大学图书馆, 聊城 252000)

[摘要] 大数据时代, 科研环境发生了变化, 数据的重要性日益凸显。数据伴随科研活动的全周期产生, 真实记录了科学研究的全过程, 是支撑和佐证科研活动的重要条件和基础。置身于海量的数据中, 如何获取、处理和利用这些数据, 成为科研人员的必备素养。英国开展数据素养教育的时间较早, 在理论和实践方面形成了完备的体系。本文从开发背景、教学对象与目标、课程内容三个方面剖析了英国伦敦大学城市学院数据素养教育课程的建设过程, 以期为我国开展数据素养教育提供借鉴。

[关键词] 高校数据; 素养教育; 课程内容

## 一、开发背景

在 E-Science 科学研究环境下, 科研范式发生了转变, 继实验科学、理论科学和计算科学之后, 第四个“数据密集型科学”研究范式出现了。其理念为“数据驱动科研发展, 科学是数据, 数据是科学”, 数据包含了重要的科研价值和使用价值。密集数据环境的形成使得大学产生的科学数据在数量、种类和速度方面都在增长, 研究人员面临一系列数据管理问题。在这种数据环境中, 对数据的获取、分析、管理与共享能力成为科研人员的必备素养。2012年, 英国皇家学会发布题为《作为开放事业的科学》的报告, 强调了数据在科研行为和交流中的重要性, 建议大学和科研机构在促进数据管理和服务方面发挥重要作用。这种专业化和规范化的数据政策无疑对于科研工作者管理科研数据的工作能力提出了更高的要求。为了契合大数据时代科研环境变化带来的挑战, 满足科研人员的数据素养教育需求, 伦敦大学城市学院开设了数据素养教育课程。

## 二、教学对象与目标

英国伦敦大学城市学院图书馆教学对象针对性较强, 包括全校的师生用户, 即本科生、研究生和科研人员。不同的教学对象不同, 教学目标也有所差异。本科生处于知识学习的初级阶段, 应注重讲解科学数据基本知识; 研究生具备了一定的研究与学习能力, 应培养科学数据管理技能; 科研人员, 具有深厚的知识基础和精湛的科研能力, 长期工作在科研一线, 需要掌握科学数据的分析、处理、存储、共享、政策等系统的数据管理知识, 以使其依托科研数据, 获取科研成果。

根据不同层次教学对象与目标的差异性, 教学课程设置有所区别。本科生主要是通识性教育, 主要介绍科学数据的基本理论和方法, 使学习者能够理解科学数据, 逐步培养数据意识。研究生与科研人员的数据素养教育模式, 以数据生命周期为指导, 以启发式教育为手段, 培养其对数据的全面理解, 从数据采集到转换应用的一系列过程以及对数据的批判性思考, 还包括针对某一具体学科领域的数据知识和技能进行教学, 以培养他们良好的数据素养。

## 三、课程内容

作为处理与图书馆信息专业相关的所有方面的数据课程, 伦敦大学城市学院的数据素养教育课程以社会技术和道德的视角, 重新定义信息技术在数据管理中的应用, 教学内容包括数据科学和数据管理基础、科学数据管理的通识性问题、科学数据管理实践等。每个课程包括两个部分: 讲解展示和实际操作。共 15 个学分, 150 个学时, 这足以确保所有学生有机会从概念和实践上获得对每个主题的理解。在一些章节中, 来自英国数字管理中心 Altmetric 等机构的客座讲师和 CILIP 提供了集体智慧。具体分为以下五个模块。

### (一) 数据概述

第一模块介绍了数据泛滥的现代现象及其对个人的影响, 信息和文件之间的关系, 计算机系统的历史发展以及计算机描述和处理数据的方式——图灵和冯诺依曼架构, 比特和字节以及编码系统, 图书馆学和信息专业的社会技术问题。确定了课程的概念框架, 并提供了对任何读者处理数据所需基本问题的解答。

### (二) 有关数据的数据

第二模块介绍了数据表单文档中的方式以及如何定义, 描述组织不同类型的文档, 从而引入元数据标准和应用程序。按照 Pomerantz (波梅兰茨) 的方法, 这会广泛地处理元数据, 对书目和 Web 资源元数据给予一些关注, 但同样关注数据集的元数据。这提供了图书馆员熟悉的元数据概念与不太熟悉的数据集上下文中的应用程序之间的链接。

### (三) 组织数据

第三模块将数据结构化为各种有组织的数据文件、平面文件、CSV 文件, 包括有关的数据库结构和标准, 包括 XML、RDF 和链接数据, 包括讨论数据管理过程, 研究数据监护的过程。对这些类型的数据文件的概念性理解和使用能力是科研人员在处理数据集合方面取得成功的基础。

### (四) 搜索数据

第四模块讨论如何在早期数据结构的基础上查找各种形式的数据。它涵盖了各种形式数据收集的搜索工具, 搜索引擎, 关系数据库系统和 SQL, 全文书目搜索系统和其他专业检索工具。它包含了大多数读者熟悉的文本检索和书目检索系统, 这些系统在更广泛的系统框架内, 具有各种检索数据的功能。

### (五) 使用数据

第五模块重点介绍数字数据和数据集, 从 Web 服务和 API (如 Twitter) 收集数据的方式, 然后进行处理、操作和分析, “数据抓取”和“数据争论”的含义; Google Sheets, Open Refine, Tags 和 Tags Explorer 等软件用于说明收集、处理、摘要和可视化。所使用的软件将随着时间的推移而变化。熟悉这些过程对于寻求成为帮助用户处理数据问题专家的图书馆员来说尤其有价值, 因为它正在成为一种广泛的数据使用形式。

## 结语

在数据密集型科学研究范式下, 伦敦大学城市学院开设的数据管理课程呈现如下特征: 注重数据能力培养的系统化, 重视信息技术和软件工具在教育课程中的应用、强调应用数据进行批判性思维的培养。目前, 我国高校还没有明确的数据素养教育课程, 我们可以参考和借鉴伦敦大学城市学院开设的数据素养教学模式, 结合本校实际, 推动我国高校数据素养教育课程的发展。□