Integrated Data Literacy Education: Theory and Practice—Based on the Postprint of Purdue University Library's "Impact: Data Science Education" Program

Authors: Yang Helin

Date: 2025-07-09T14:53:06+00:00

Abstract

How to integrate data literacy education into professional courses is one of the current research hotspots in the field of data literacy education in university libraries. The Purdue University Library in the United States has proposed an integrated data literacy education model that innovatively adopts an approach of supporting faculty development and optimizing course assignments rather than direct instruction to carry out related work. Through literature and web-based research methods, this article examines the implementation framework, specific practices, theoretical foundations, and advantages of the library's "Impact: Data Science Education" project, providing practical and valuable references for university libraries in China in optimizing theoretical support for data literacy education, building positive faculty-librarian relationships, strengthening campus collaboration, and securing industry support.

Full Text

Abstract

How to integrate data literacy education into disciplinary curricula is a current research focus in university library data literacy education. Purdue University Library has proposed an integrated data literacy education model that innovatively employs faculty development and curriculum assignment optimization rather than direct instruction. This paper examines the implementation framework, specific practices, theoretical foundations, and advantages of the library's "IMPACT: Data Science Education" (IDSE) project through literature review and web-based research methods. The study offers practically valuable references for Chinese university libraries in strengthening theoretical support

for data literacy education, building effective faculty-librarian partnerships, enhancing campus-wide collaboration, and securing industry support.

Keywords: University library; Data literacy; Data-informed learning

---

1. Overview of Purdue University Library's Data Services

Data literacy is regarded as one of the most important new competencies of the 21st century. The OECD defines it as a core foundation for learners in the data-intensive era. China's *Scientific Data Management Measures* and *Action Outline for Improving National Digital Literacy and Skills* emphasize strengthening scientific data workforce development and exploring new data-driven research paradigms. University libraries, as planners and key drivers of information literacy education, have extended data literacy from information literacy and launched a series of data literacy education initiatives.

Purdue University Library's data services have matured through years of exploration, now offering a comprehensive suite of services to campus (Table 1). Key services include the Purdue University Research Repository (PURR) for dataset publication and archiving, Purdue e-Pubs for institutional publications, Data Management Toolkit for data management planning, and the Distributed Data Curation Center (D2C2). The library also offers multiple data-related courses (Table 2), continuously accumulating data teaching experience that laid the foundation for the IDSE project.

Table 1: Purdue University Library's Main Data Services

Service	Description
Purdue University Research Repository (PURR)	Online collaborative workspace and platform for publishing and archiving datasets
Purdue e-Pubs	Institutional repository for publications and white papers
Data Repositories Bibliography (Databib)	Tool for identifying and locating research data repositories
re3data.org	International registry of research data repositories
Data Management Toolkit	Collection of open-access data management tools
Distributed Curation Center (D2C2)	Virtual center for data curator career development and research

Table 2: Selected Data-Related Courses at Purdue University Library

Course	Description
Quantitative Research Data Management	Covers dataset organization, versioning, quality control, and visualization analysis
Intro to Data Lifecycle Management	Systematic study of data creation, preservation, and ethics
Data Science & Society: Legal and Social Impact (ELSI)	Explores ethical, legal, and social issues in data science
Understanding Your Research Data	Teaches data management and visualization principles for undergraduates
Geo Data Science with ArcGIS Python	Provides hands-on data operation environment in ArcGIS Python scripting
Intro to Information Studies	Critical understanding of data, intellectual property in data-intensive research

2. Implementation Framework of the IDSE Project

Purdue University Library operates on the principle of supporting faculty development and curriculum redesign rather than direct teaching. Guided by educational theory and the latest research, the library facilitates faculty in transforming undergraduate courses through Faculty Learning Communities (FLCs), incorporating active learning and student-centered pedagogies. This deep collaboration strengthens the teaching and research support system while enhancing students' critical thinking, complex reasoning, and scientific writing skills.

The implementation framework comprises three core components:

Faculty Development through IMPACT: Purdue's "Instruction Matters: Purdue Academic Course Transformation" (IMPACT) program, established in 2011, forms FLCs to support course redesign. The library, as a key IMPACT partner, helps faculty integrate data assignments into disciplinary courses. Over 400 courses have been reformed, with faculty reporting significantly improved teaching practices, student engagement, and satisfaction.

Student Data Literacy through Data Community: The Integrative Data Science Initiative (IDSI), launched by the Provost's Office, builds a campus-wide data science ecosystem. The Data Community connects students with data-related courses, research projects, and scholarships. IDSE project participants must commit to integrating data assignments into their courses, excluding courses already in data science programs.

Data Assignment Development: IDSE supports faculty in designing data-centric assignments ranging from basic data literacy to computational methods for analyzing diverse datasets. The library introduces faculty to data resources and services across campus, with dedicated funding supporting promising teaching initiatives.

3. Project Practices

Addressing Faculty Needs: As data-intensive research grows, faculty face challenges in designing data instruction. IDSE connects them with teaching centers and data science support services. The library assists in finding appropriate datasets, evaluating data quality, and providing technical tools. Faculty receive guidance on data policies, high-performance computing, and tools like Python and Jupyter Notebook through workshops and individual consultations.

Collaborative Assignment Design: Assignment development begins with collective brainstorming sessions where faculty draft objectives and design ideas. The library facilitates discussions on data literacy theory, recommends resources, and connects faculty with data librarians and subject specialists. Experienced faculty share best practices—for example, a biology professor integrating research projects into foundational courses, and a GIS expert demonstrating data assignment design in disciplinary contexts.

Data-Informed Learning Model: The library adopts Christine Bruce's "Informed Learning" concept, proposing "Data-Informed Learning" as an innovative approach. This model positions data use learning alongside disciplinary knowledge and professional practice, emphasizing creative and reflective data utilization in authentic contexts. Key principles include:

- **Building on Prior Knowledge:** Students first master disciplinary foundations before engaging with data operations and reflection, avoiding pure discovery learning risks.
- **Synchronizing Data and Professional Learning:** Data literacy and disciplinary knowledge develop spirally through bidirectional empowerment mechanisms.
- **Connecting to Real-World Cognition:** Students model real-world phenomena to develop engineering solutions, as in environmental engineering courses using domain-specific data.
- **Leveraging Educational Variation Theory:** Deepening data understanding through comparison of similarities and differences.

4. Theoretical Analysis

From "One-Shot" to "Train-the-Trainer": Traditional one-shot library instruction lacks scalability and sustainability, underutilizing librarians' educational potential. IDSE's "train-the-trainer" model empowers faculty to become instructional executors, enabling deep curriculum integration and sustainable scaling. Librarians transition from peripheral executors to core creators of educational value.

Educational Development-Based Team Building: Effective data literacy education requires cross-institutional curriculum development teams comprising faculty, librarians, teaching administrators, and technical staff. Team members must share goals, communicate respectfully, and collaborate under mutual trust.

This approach helps librarians consolidate their teaching status and achieve integration with professional courses.

Data-Informed Learning Design: Modern educational psychology reveals that learning occurs through interpersonal interaction before internalization. Librarians serve as consultants and technical supporters, helping faculty innovate pedagogically. The model respects faculty leadership while positioning librarians as developers who assist in goal-setting, task design, and resource integration.

5. Project Advantages

Theory-Practice Integration: The project demonstrates that theory not only explains reality but also guides practice. Purdue's approach—combining data-informed learning theory with faculty development, stakeholder collaboration, and continuous assessment—provides a replicable model. Faculty value practical, evidence-based guidance over abstract concepts.

Respecting Faculty Leadership: Unlike approaches that push pre-designed content, IDSE positions faculty as data leaders in teaching. This respects disciplinary expertise and teaching autonomy while addressing faculty concerns about data quality, tool selection, and learning objectives. Even faculty with weak data literacy receive tailored support without compromising their instructional authority.

Stakeholder Collaboration: The project successfully engages diverse stakeholders—faculty, librarians, teaching centers, computing services, and industry associations like IMLS and ALA. This multi-party collaboration creates synergies, with the library serving as a bridge connecting teaching, research, and data services.

6. Implications for Chinese University Libraries

Strengthening Theoretical Support: Chinese libraries should deepen theoretical research on integrated data literacy, particularly data-informed learning models. Theoretical frameworks help librarians design effective strategies and gain faculty trust.

Building Faculty-Librarian Partnerships: Libraries must transition from passive supporters to active co-creators. By participating in curriculum design as developers, connectors, and facilitators, librarians can embed data literacy into disciplinary systems and achieve mutual empowerment.

Leveraging Institutional Initiatives: Integrated data literacy education should align with university digital transformation strategies. For example, China's national digital literacy initiatives and the "Digital China" strategy provide policy windows for libraries to integrate data services into institutional reforms.

Securing Industry Support: Professional associations play crucial roles. The Chinese University Library Research Data Management Working Group, Wuhan University's National Digital Literacy Training Base, and Fudan Library's Open Data Innovation Competition demonstrate how industry collaboration can amplify impact. Libraries should actively participate in working groups, share best practices through workshops, and engage with international organizations.

7. Conclusion

Purdue University Library's integrated data literacy education model offers an innovative approach that prioritizes faculty development over direct instruction. Through the IMPACT: Data Science Education project, the library has created a sustainable, scalable framework that respects disciplinary expertise while embedding data literacy across the curriculum. The model's emphasis on theoretical grounding, collaborative partnership, and stakeholder engagement provides valuable guidance for Chinese university libraries. As China advances its national digital literacy strategy, libraries should seize opportunities to become core partners in institutional digital transformation, contributing wisdom and strength to improving national digital literacy and skills.

References

[1] OECD. Core foundations for 2030[EB/OL]. [2024-11-26]. https://www.oecd.org/content/dam/oecd/en/abou 2040/concept-notes/Core_{{{Foundations}}*{{for}}*{{{2030}}*{{concept}}*}{note}.pdf.

[2] General Office of the State Council. Notice on issuing the Scientific Data Management Measures[EB/OL]. [2024-10-07]. https://www.gov.cn/gongbao/content/2018/content_{5283177}

[3] China Cyberspace Administration. Action outline for improving national digital literacy and skills[EB/OL]. [2024-10-17]. https://www.cac.gov.cn/2021-11/05/c_{1637708867754305}.htm.

[4] Ebijuwa A S, Tella A. Integrating data literacy into information literacy programmes: any difference?[J]. International Journal of Library and Information Science Studies, 2023(6): 33-45.

[5] Matuk C, Vacca R, Amato A. Promoting students' informal inferential reasoning through arts-integrated literacy education[J]. Information and Learning Sciences, 2024(3/4): 163-189.

[6] Yatcilla J K, Maybee C, McGowan B. Teaching data in the social sciences: the Purdue report[EB/OL]. [2024-09-04]. https://docs.lib.purdue.edu/libreports/6/.

[7] Hammons J. Academic librarians as teachers and faculty developers: exploring the potential of a "teach the teachers" model[J]. College and Research Libraries, 2024(3): 379-398.

[8] Bruce C, Hughes H. Informed learning: a pedagogical construct attending simultaneously to information use and learning[J]. Library and Information Science Research, 2010(4): A2-A8.

[9] Maybee C, Zilinski L D. Data-informed learning: a phase framework for higher education[J]. Proceedings of the Association for Information Science and Technology, 2015(1): 1-4.

[10] Association of College and Research Libraries. Framework for information literacy for higher education[EB/OL]. [2024-10-04]. https://www.ala.org/sites/default/files/acrl/content/issues

[11] Purdue University Library. Research data[EB/OL]. [2024-11-15]. https://www.lib.purdue.edu/researchdata.

[12] Purdue University School of Information Studies. Course listing[EB/OL]. [2024-10-15]. https://infostudies.purdue.edu/course-listing/.

[13] Purdue Evaluation and Learning Research Center. IMPACT report summer 2023: a report from the IMPACT evaluation team[EB/OL]. [2024-09-29]. https://docs.lib.purdue.edu/impactreps/15/.

[14] Purdue IMPACT. About IMPACT[EB/OL]. [2024-09-01]. https://www.purdue.edu/impact/about/.

[15] Purdue University. What is data science?[EB/OL]. [2024-11-19]. https://www.science.purdue.edu/data-science/index.html.

[16] Purdue University. The Data Mine[EB/OL]. [2024-10-06]. https://datamine.purdue.edu/.

[17] Purdue University. Office of the Provost announces request for proposals to build integrative data science education ecosystem[EB/OL]. [2024-10-06]. https://www.purdue.edu/newsroom/purduetoday/releases/2018/Q2/office-of-the-provost-announces-request-for-proposals-to-build-integrative-data-science-education-ecosystem.html.

[18] Purdue University. IMPACT data science education[EB/OL]. [2024-10-23]. https://sites.lib.purdue.edu/dse/.

[19] National Academies of Sciences, Engineering, and Medicine. Data science for undergraduates: opportunities and options[EB/OL]. [2024-09-03]. https://nap.nationalacademies.org/read/25104/chapter/4#22.

[20] Purdue Rosen Center for Advanced Computing. Collaboration partners[EB/OL]. [2024-10-03]. https://www.rcac.purdue.edu/about/collaboration.

[21] Bauder M, Maybee C, Fundator R. Academic librarians as informed learning developers: a phenomenographic study[J]. Communications in Information Literacy, 2019(2): 184-204.

[22] Henderson J, D'Ignazio C. Lessons learned from a two-year experience teaching data literacy to faculty[EB/OL]. [2024-11-29]. http://docs.lib.purdue.edu/iatul2010/conf/day2/5.

[23] Timmermans J A, Berka C S. A comparative analysis of data literacy competency frameworks[EB/OL]. [2024-09-25]. https://iopn.library.illinois.edu/journals/aliseacp/article/view/1046/

[24] Kim J, Corry M. Data literacy training for education professionals[J]. Journal of Research on Technology in Education, 2021(2): 232-244.

[25] Flierl M, Maybee C, Fundator R. Academic librarians as informed learning developers[M]//Informed learning applications: insights from practice. Emerald Publishing Limited, 2019: 81-94.

[26] Maybee C, Bruce S, Lupton M. Informed learning design: identifying threshold concepts for information literacy developers[J]. International Journal for Academic Development, 2014(4): 305-317.

[27] Pingree S. Creating future faculty development: learning from the past, understanding the present[J]. Review of Higher Education, 2006(1): 85-87.

[28] Wang L. Information literacy integration model and application in higher education[J]. Reference Services Review, 2011(4): 703-720.

[29] Fundator R. Academic librarians as informed learning developers: a phenomenographic study[J]. Communications in Information Literacy, 2019(2): 184-204.

[30] Adams J, Maybee C, Fundator R. Academic librarians' experiences as informed learning developers[J]. Journal of Research in Innovative Teaching & Learning, 2021(2): 232-244.

[31] Hogenboom K, Holler P, Carissa P. Data! Data! Data! Partnering with instructors to teach data literacy[EB/OL]. [2025-02-27]. https://alair.ala.org/items/aa5120b0-fcc1-466f-ac0d-248bf21bc815.

[32] Qin L. Research on university library data literacy education models[J]. University Library Journal, 2022(7): 10-14, 28.

[33] Carlson J, Johnston L. Data information literacy: librarians, data, and the education of a new generation of researchers[M]. Indiana: Purdue University Press, 2015: 1-8.

[34] Henderson J, D'Ignazio C. Lessons learned from a two-year experience teaching data literacy to faculty[EB/OL]. [2024-11-29]. http://docs.lib.purdue.edu/iatul2010/conf/day2/5.

[35] Kim J, Corry M. Data literacy training for education professionals[J]. Journal of Research on Technology in Education, 2021(2): 232-244.

[36] Carlson J, Johnston L, Westra B. Developing an understanding of data literacy education[J]. Journal of Education for Library and Information Science, 2013(2): 79-88.

[37] Ministry of Education. Empowering high-quality education development with innovation elements[EB/OL]. [2024-09-17]. http://www.moe.gov.cn/jyb_{xwfb}/s5148/202405/t20240508

[38] Yang X C. The reality dilemma and development path of libraries in data literacy education[J]. Jiangsu Science and Technology Information, 2024(3): 72-76.

[39] Adams J, Maybee C, Fundator R. Academic librarians' experiences as informed learning developers[J]. Journal of Research in Innovative Teaching & Learning, 2021(2): 232-244.

[40] Maybee C, Bruce S, Lupton M. Informed learning design: identifying threshold concepts for information literacy developers[J]. International Journal for Academic Development, 2014(4): 305-317.

[41] China Education Daily. Theory and its dialectical relationship with practice[N]. 2009-11-24(11).

[42] Flierl M, Maybee C, Fundator R. Academic librarians as informed learning developers[M]//Informed learning applications: insights from practice. Emerald Publishing Limited, 2019: 81-94.

[43] Carlson J, Johnston L. Data information literacy: librarians, data, and the education of a new generation of researchers[M]. Indiana: Purdue University Press, 2015: 1-8.

[44] Shanghai Library Society. Introduction to the Chinese University Research Data Management Advancement Working Group[EB/OL]. [2024-12-04]. http://society.library.sh.cn/node/2611.

[45] Wuhan University School of Information Management. National digital literacy capacity enhancement special topic[EB/OL]. [2024-12-05]. https://sim.whu.edu.cn/info/5341/93202.htm.

[46] Shanghai Jiao Tong University Library. 2024 China University Research Data Management Advancement Working Group meeting held[EB/OL]. [2024-09-17]. https://www.lib.sjtu.edu.cn/f/content/detail.shtml?id=7667&lang=zh-cn.

[47] Yang X C. Exploring data literacy education models in social science data platforms[J]. Library and Information Service, 2024(11): 82-91.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.