



# The Development of an Undergraduate Data Curriculum: A Model for Maximizing Curricular Partnerships and Opportunities

Angela P. Murillo<sup>(✉)</sup>  and Kyle M. L. Jones

Indiana University – Purdue University Indianapolis,  
Indianapolis, IN 46202, USA  
apmurill@iu.edu, kmlj@iupui.edu

**Abstract.** The article provides the motivations and foundations for creating an interdisciplinary program between a Library and Information Science department and a Human-Centered Computing department. The program focuses on data studies and data science concepts, issues, and skill sets. In the paper, we analyze trends in Library and Information Science curricula, the emergence of data-related Library and Information Science curricula, and interdisciplinary data-related curricula. Then, we describe the development of the undergraduate data curriculum and provide the institutional context; discuss collaboration and resource optimization; provide justifications and workforce alignment; and detail the minor, major, and graduate opportunities. Finally, we argue that the proposed program holds the potential to model interdisciplinary, holistic data-centered curriculum development by complimenting Library and Information Science traditions (e.g., information organization, access, and ethics) with scholarly work in data science, specifically data visualization and analytics. There is a significant opportunity for Library and Information Science to add value to data science and analytics curricula, and vice versa.

**Keywords:** Curriculum · Undergraduate education · Data studies  
Data science · Data analytics · Library and Information Science  
iSchool

## 1 Introduction

Our data-intensive world requires individuals to build data-specific skill sets to participate in the workforce and society. In recent years, we have seen the value of data skills increase in academia, business, and government. Additionally, we have seen the emergence of data-focused programs to provide students the skills needed to succeed in nearly all domains and disciplines. The current situation for employers is dire with an estimated dearth of nearly 200,000 laborers needed with “deep analytical skills” and 1.5 million employees who lack the knowledge to make effective decisions using data-driven insights [1].

As a discipline, Library and Information Science (LIS) has developed theories, practical frameworks, and critical perspectives centered around data. And while modern LIS curricula have largely concentrated on graduate education, there is

potential to translate data-focused curricula for the education of undergraduate students. Data organization and representation, data curation and management, and data policy and ethics are all perspectives and skill sets that would benefit undergraduates. With these opportunities in mind, the Department of Library and Information Science (DLIS) at Indiana University–Purdue University Indianapolis (IUPUI), began its Data Studies Minor program and developed the Applied Data and Information Science (ADIS) undergraduate major in 2017.

The department’s motivation was to provide students across the campus access to a data-focused curriculum that would provide students skills necessary to enter the data workforce and analyze data practices. This minor and major provides students the opportunity to learn how to manage data, curate data, as well as understand the policy and governance issues associated with data work—all things the workforce increasingly needs as it seeks to maximize value from big and small data practices alike. Perhaps more importantly, it provides students a set of skills and critical perspectives that are valuable for participating in our information society. As one of three interdisciplinary departments in the School of Informatics and Computing (SoIC), DLIS was well positioned to accomplish these goals.

For this paper, we will use the term “data-focused” to discuss in general data-related curricula. Additionally, we will use the term “data studies” to discuss the original vision for the data studies minor. Data Studies “combines data acquisition, management, analysis, and use of data...with an understanding of the nature of data and its broader implications for society” [2]. Lastly, we will use the term “data science” to describe curriculum “more associated with data analytics and computer science which implies a specific set of content that is more technical in nature” [2].

Moving forward, we analyze current LIS curricula and explore data-focused curricula. Next, we describe the development of the undergraduate minor and major program developed at IUPUI, and we cover the institutional and labor contexts that informed the program development. Finally, we argue that the proposed program holds the potential to model interdisciplinary, holistic curriculum development at the undergraduate level.

## 2 Undergraduate Curricula

### 2.1 Trends in Library and Information Science

Library and Information Science (LIS) educators have historically focused on graduate education; however, in recent years we are seeing a resurgence of undergraduate curriculum development with a focus on preparing students for careers in information industries. Modern undergraduate program development began in the 1990s. Florida State University’s undergraduate program, for instance, began in 1996 with 91 students and by 2002 had over 600 enrollments. Undergraduate enrollment at Syracuse and Rutgers saw similar upward enrollment trends from the late 1990s to the early 2000s [3]. Non-United States undergraduate LIS programs include Australia [4], Latin America and the Caribbean [5], and Europe where there are more undergraduate programs available than graduate programs [6].

In the early 1990s, the demand for students who understood emerging information and communication technologies in a variety of contexts motivated many institutions to create undergraduate programs. For example, the Rutgers and University of Washington undergraduate programs focused on informatics, the University of Wisconsin-Milwaukee and Emporia State program focused on information resources, and the Pittsburgh, SUNY, and Southern Mississippi programs focused on information science [3]. The LIS community continued to create new undergraduate programs throughout the 2000s with the University of North Carolina in 2003 [7] and University of Texas-Austin, which created a minor in 2005 [8]. All of this activity points towards changes in the scope of LIS education outside of graduate programs, indicating that there is a path for curricular influence by LIS educators in undergraduate programs when new opportunities emerge.

## 2.2 The Emergence of Library and Information Science Data Curricula

Outside of creating undergraduate programs, LIS educators have created certificate programs and specializations to consider the many facets of data labor. From 2006–2010, UNC-Chapel Hill, the University of Illinois at Urbana Champaign, University of Arizona, and the University of Michigan at Ann Arbor developed Digital Curation Education certificate programs [9, 10]. These programs focus on skills needed for digital or data curation including digital presentation, curation, records management, digital collections, and data manipulation. Other universities developed similar programs with specific subject foci, including eScience at Syracuse and Cultural Heritage at Pratt Institute [9].

The LIS community has seen a rise in curriculum development with other data related or technology related skill sets, as well as domain-specific skill sets. These included addressing the curriculum gap between LIS education and intelligence analysis. As described in Jin and Bouthillier [11] and Wu [12], while LIS programs provide education in computer skills for information collections, geographic information systems, and information security, there is a lack of programming in science and technology intelligence. Additionally, there have been several efforts to develop LIS curriculum and eScience and/or scientific data management; however while this curriculum is in place to educate students and there is demand for experts, there is still yet to be a unified title for the scientific data specialization [13]. Domain science educators have indicated that the majority of the work in data literacy has focused on graduate students, while there is a real need for similar data literacy instruction to include undergraduate students [14].

As described above, LIS curricula already include courses that are specific for data professionals. In a study by Varvel Jr. et al. (2012) 475 courses in 158 programs at 55 schools contained at least some aspects of data education [15]. These courses included data-centric courses related to data curation, data management, or data science topics, data inclusive courses related to e-science or e-research, digital courses including digital libraries, collections, and preservation; they also included traditional LIS courses, which introduced important topics to be developed further

in data inclusive or data-centric courses [15]. There has been a steady growth in data related curriculum in LIS schools to educate the future data workforce and to provide students the skills needed to work in data-intensive environments. Tonta provides an overview of the major developments in LIS education from 1887–1963 (the first period), 1994–1993 (the second period), and 1994–the present (the third period). As described by Tonta, the third period is shifting curriculum to information science, information systems, technologies, as well as overlapping some with computer science in certain topics such as information retrieval, social informatics, and infometrics [16]. All of this activity points towards an increasing growth of data related curricula in LIS graduate programs.

### 2.3 Interdisciplinary Opportunities

While the LIS community has been developing data-focused curricula, and LIS research has become interdisciplinary, there has also been a significant development in data science undergraduate curricula. Studies have analyzed the differences between undergraduate data analytics programs and undergraduate data studies programs with data science programs. Data science often requires more mathematical and programming courses than data analytics and data studies programs. Data analytics programs emphasize the use of case studies and evaluations of tools, while data science programs emphasize the implementation of tools, techniques, and visualization strategies [17]. There has also been a history of establishing undergraduate data science programs through interdisciplinary collaborations, resulting in joins of curricula from multiple disciplines. As shown in Anderson, 2014, three undergraduate programs often inform interdisciplinary efforts in data science: (1) computer science and mathematics, (2) computer science, business informatics, and statistics, and (3) computer science, mathematics, and economics [18]. The study of data science programs has also indicated that data science mirrors the interdisciplinary nature of data, and that curricula should balance theory and concepts, tools and techniques, and should include contributions from computer science, artificial intelligence, mathematics, statistics, data mining, communication, and discussion of social and ethical issues [19]. Other studies indicate the need to include specific skill sets including data visualization, data manipulation/data wrangling, computational statistics, and machine learning in data science curricula [20]. Lastly, the six main subject areas of data science are data description and curation, mathematical foundations, computational thinking, statistical thinking, data modeling, and communication, reproducibility, and ethics [21].

While LIS has traditionally taught subjects important to data science curriculum including data description, curation, representation, management, policy, and ethics—among other things—certain topics are more typically taught in other departments such as data manipulation, visualization, and analytics. With this in mind, DLIS along with DHCC at IUPUI can provide the full spectrum of data-related curriculum, providing SoIC the unique opportunity to create a data-focused curriculum for undergraduates.

## 3 Developing an Undergraduate Data Curriculum

### 3.1 Institutional Context

The IUPUI Department of Library and Information Science (DLIS) is within the School of Informatics and Computing (SoIC), which also contains the Human-Centered Computing (DHCC) and BioHealth Informatics departments (DBHI) [22]. SOIC's mission is to excel in education, research, and civic engagement in the field of informatics, an integrative discipline which advances knowledge in: (1) computing, information, and media technologies, (2) the implication those technologies have for individuals and society; and, (3) their application to any field of study adapting to the challenges of the Information Age and fosters a broad and interdisciplinary view [23]. It is with this mission in mind DLIS partnered with DHCC to create the programs described below.

### 3.2 Disciplinary Collaborations and Resource Optimization

DLIS saw an opportunity to develop a complete undergraduate curriculum as the structure of SoIC provided the unique opportunity for DLIS to leverage already existing resources and make strategic partnerships with research and teaching faculty in DHCC. The collaboration between DLIS and DHCC focused on providing a holistic approach to a data-focused curriculum. Together, they proposed a major with two specializations. The program is titled Applied Data and Information Science (ADIS), and its specializations include Applied Data Science (ADS) and Information Science (IS). The departments worked together incorporating courses from both departments to create the major curriculum. Additionally, the departments worked together to update already existing courses for the major and created new courses as necessary.

The DHCC faculty provides courses in informatics, visualization, cloud computing, and data analytics and DLIS faculty provides courses in data representation, organization, preservation, curation, policy and ethics, and socio-technical analysis. Since DLIS and HCC have collaborated on this program together, we are financially benefiting each other but not competing with each other. At SOIC, the instructor of record's home department receives the funds from the courses taught. Additionally, technology and staff resources are shared at SoIC, providing shared support for the program regardless of departmental affiliation.

### 3.3 Justifications and Workforce Alignment

#### Data Labor

Data labor skill sets have been well-documented in data science and data lifecycle workflows such as the OSEMN taxonomy of data science which provides the types of data skills needed by data scientists, including obtaining data from multiple sources, scrubbing data, exploring data using visualization techniques, modeling data, and interpreting data [24]. Additionally, the DataONE lifecycle has provided an understanding of the data lifecycle, as well as the various skills needed for successful

management, curation, and preservation of data. These skill sets include creating data plans, collecting data, assuring the quality of data, describing data, understanding data preservation, discovering data, integrating data from multiple data sources, and lastly understanding various data analysis techniques [25]. While each of these models describes the skill sets needed for data labor and the activities involved in data labor, there is some overlap. Additionally, nearly every field needs these skill sets, and the proposed program provides students a unique opportunity by ensuring that students can gain experience in the entire spectrum of data labor skill sets.

### **Indianapolis and Indiana**

This new program is serving the needs of the state and community. As a public university, our role is to provide education that is beneficial to our community. Indiana and Indianapolis have been working to promote the technology industry, and by training students in data, we are assisting in developing a data-savvy workforce for Indiana.

Trends related to high-tech labor highlight Indiana and Indianapolis as a thriving market, a standout in fact when compared to other so-called “flyover” regions of the country [26]. According to Moody’s data, as analyzed by the Brookings Institute, the Indianapolis area created around 5,000 high-tech jobs between 2013 and 2015, which marked a 13.9% increase and even beat San Francisco’s growth rate [27]. In July 2016, former governor Mike Pence pledged a total of \$1 billion “to advance innovation and entrepreneurship in Indiana,” out of which the governor’s office directed \$300 million to the 21 Fund for research and technology and \$100 million to “further advance innovation and entrepreneurship education” [28]. Even after a transition in the governor’s office, the funding commitments remain, and the state is on track to attract innovative talent and educate the next leaders in information technology. Investments by the state have started to pay off. According to the Central Indiana Corporate Partnership report, “since 2007, 12 tech community companies in Central Indiana have either been acquired or have gone public, generating \$4.5 billion in market value and creating more than 3,700 Indiana jobs, demonstrating the momentum of the tech sector and the continued need for top talent” [29]. Indiana higher education institutions, including IUPUI, contribute to the labor workforce of these companies and the economic programs the state deeply supports.

### **3.4 The Minor**

The initial program development began with the creation of a 15 credit hour Data Studies minor [30]. The minor was created to provide a humanistic approach to data—both big and small—and to train students on data management and curation, and to introduce students to data analysis, manipulation, organization and representation, and policy. The minor was created to make students more marketable in any field. We created several new undergraduate courses for this minor and pulled from already existing LIS areas of study, including: data organization and representation, data curation and management, and data policy and governance. The foundational course, Foundations in Data Studies, provides students an overview of the many topics of the minor and prepares the students for the program. From this program, students will be

able to apply principles of representation and organization, understand and apply data curation processes, and develop data policy. Additionally, a minor focused on applied data science was created and focuses on data analytics and technical aspects of working with data [31]. These minors are currently available to all students at IUPUI and stand alone as complimentary curricular additions to any student's major.

### 3.5 The Major

After developing the Data Studies minor, DLIS and DHCC together proposed the 120-credit hour ADIS undergraduate major. This major provides students an additional avenue to learn more extensive data skill sets, as well as prepare them for graduate studies in data-focused or LIS programs if they so choose. ADIS students earn a bachelor of science degree with one of two specializations: Applied Data Science (ADS) and Information Science (IS). The purpose of having these two specializations is to encompass the entire spectrum of data labor skill sets. The ADS specialization focuses on aspects of data labor involved with data analysis, visualization, and the more technical aspects of data labor, which is a good fit for students comfortable in mathematics and computing. The IS specialization focuses on data curation, organization, and management, which is a good fit for students who are less comfortable in mathematics, but still interested in learning the technical skills needed to work with data. All students have to take courses in the other specialty. For example, all ADS students will take data curation and all IS students will take data visualization, as well as policy and ethics courses to provide them holistic knowledge of data skills.

Additionally, students in both specializations take courses regarding societal implications of data labor including courses related to ethics, policy, surveillance, and privacy. Furthermore, we are creating a wide variety of elective courses for students including courses focused on specific types of data (business data, scientific data, social science data), as well as specific relevant topics (data and society, data archives, data of social media). These courses will help provide students a more refined study of data labor of interest to them or their future employment. Students will also be involved in internships to provide them real-world experience and will have a capstone experience as they complete their studies.

### 3.6 Graduate Opportunities

The program holds the excellent potential to bridge students into graduate programs at IUPUI, especially within SoIC. The workforce data supports our argument that ADIS students will be employable with respectable salaries, but we do not deny that the same data also shows even greater potential to make marked salary jumps if ADIS students pursue a graduate degree. Should students seek a master or doctoral-level degree, a background in ADIS from either proposed specialization will surely benefit students. SoIC currently offers the following graduate degree options that align with the ADIS curriculum, many of which provide flexible learning arrangements:

- Library and Information Science (M.L.S.)
- Bioinformatics (M.S. and Ph.D.)

- Applied Data Science (M.S.)
- Sports Analytics (M.S.)
- Data Science (Ph.D.)

Students coming from a health background who pursue a minor in ADIS will also be more prepared for these SoIC programs:

- Bioinformatics (M.S.)
- Health Informatics (M.S.)
- Health/Biomedical Informatics (Ph.D.)

All of the above programs require a statistical foundation, technical skill sets, and sensitivity to the social, legal, and ethical implications related to implementing data-driven technologies and building socio-technical infrastructures—all of which ADIS provides.

Given how the proposed ADIS program will prepare students to pursue related graduate degrees, there existed an excellent opportunity to develop the program into a 4 + 1 model (four years to earn a bachelor degree, one year to earn a master degree). We developed two 4 + 1 options for students to enter into the Applied Data Science or Library and Information Science SoIC master programs. That said, we are currently pursuing opportunities to work with other departments within and outside of SoIC to co-develop more 4 + 1 degree options. The creation of this undergraduate program completed the educational pipeline into SoIC's master and doctoral-level programs focused on data, as well as into the master of library science program.

## 4 Conclusion

Library and Information Science programs are already well-established in educating in specific areas of data labor skill sets, including data curation, preservations, representation, organization, and research data management. While these skill sets do not encompass the entire spectrum of data labor, they do encompass a rather substantive portion. With this in mind, the Department of Library and Information Science (DLIS) at IUPUI was in the unique and promising position to collaborate with the Department of Human-Centered Computing (DHCC) to create a data program that encompasses the majority of data skills needed in today's data-intensive world.

With the ever-changing and transforming world of technology and data, we envision these types of curricular collaborations to become more relevant and necessary to ensure that our students have the education and skills they need to enter the emerging data-driven workforce. As discussed, the LIS community already has a well-established history of teaching data and information literacy-related courses and the expertise in specific data topics. However, with other important data topics such as data analytics, data mining, and visualization; it is sensible to partner with related disciplines. We envision our model for program development as an innovative and inspiring way to conceptualize LIS undergraduate education.



## References

1. Brown, M.S.: What analytics talent shortage? How to get and keep the talent you need. *Forbes*, 27 June 2016. <https://www.forbes.com/sites/metabrown/2016/06/27/what-analytics-talent-shortage-how-to-get-and-keep-the-talent-you-need/>. Accessed 18 Sept 2017
2. UNC Chapel Hill Faculty Working Group on Data Studies Curriculum: Developing data-literate students acquiring, managing, analyzing, and using data in societal context. Report and Recommendations (2014). <http://innovate.unc.edu/wp-content/uploads/2016/10/Data-Studies.pdf>
3. McInerney, C., Delay, A., Vandergrift, K.E.: Broadening our reach: LIS education for undergraduates. *Am. Libr.* **33**(2), 40–43 (2002). [https://www.jstor.org/stable/25646224?seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/25646224?seq=1#page_scan_tab_contents)
4. Sanders, R.: Current demand and future need for undergraduate LIS education. *Aust. Libr. J.* **57**(2), 102–127 (2008)
5. Gallardo, A.R.: Library education in Latin America and the Caribbean. *New Libr. World* **108** (1/2), 40–54 (2007)
6. Borrego, A.: Library and information education in Europe: an overview. *BiD: Textos universitaris de biblioteconomia i documentació* 35 (2015)
7. UNC Bachelor of Science in Information Science (BSIS). <https://sils.unc.edu/programs/undergraduate/bsis>. Accessed 1 Sept 2017
8. Roy, L., Simmons, R.N.: Tradition and transition: the journey of an iSchool deep in the heart of Texas. *J. Libr. Inf. Technol.* **37**(1), 3–8 (2017)
9. Fulton, B., Botticelli, P., Bradley, J.: DigIn: a hands-on approach to a digital curation curriculum for professional development. *J. Educ. Libr. Inf. Sci.* **52**(2), 95–109 (2011)
10. Murillo, A.P., Barnes, H.L., Poole, A.H., Uckum, T.: Digital curation education programs in the United States. In: DCC 7th International Digital Curation Centre Conference, Bristol, England (2011)
11. Jin, T., Bouthillier, F.: The integration of intelligence analysis into LIS education. *J. Educ. Libr. Inf. Sci.* **53**(2), 130–148 (2012)
12. Wu, Y.: A preliminary study on the curriculum overlap and gap between LIS education and intelligence education. *J. Educ. Libr. Inf. Sci.* **54**(4), 270–285 (2013)
13. Li, S., Xiaozhe, Z., Wenming, X., Weining, G.: The cultivation of scientific data specialists: development of LIS education oriented to e-science service requirements. *Libr. Hi Tech* **31** (4), 700–724 (2013)
14. Reisner, B.A., Vaughan, K.T.L., Shorish, Y.L.: Making data management accessible in the undergraduate chemistry curriculum. *J. Chem. Educ.* **91**, 1943–1946 (2014)
15. Varvel Jr., V.E., Bammeril, E.J., Palmer, C.L.: Education for data professionals: a study of current courses and programs. In: *Proceedings of iConference 2012*, pp. 527–529. ACM, Toronto (2012)
16. Tonta, Y.: Developments in education for information: will “data” trigger the next wave of curriculum in LIS schools? *Pak. J. Inf. Manag. Libr.* **17**, 2–4 (2016)
17. Aasheim, C.L., Williams, S., Rutner, P., Gardiner, A.: Data analytics vs. data science: a study of similarities and differences in undergraduate programs based on course descriptions. *J. Inf. Syst. Educ.* **26**(2), 103–115 (2015)
18. Anderson, P., McGoufee, J., Uminsky, D.: Data science as an undergraduate degree. In: *Proceedings of SIGCSE 2014*, pp. 705–706. ACM, Atlanta (2014)
19. Anderson, P., Bowring, J., McCauley, R., Pothering, G., Starr, C.: An undergraduate degree in data science: curriculum and a decade of implementation experience. In *Proceedings of SIGCSE 2014*, pp. 145–150. ACM, Atlanta (2014)

20. Baumer, B.: A data science course for undergraduates: thinking with data. *Am. Stat.* **69**(4), 334–342 (2015)
21. De Veaux, R.D., Agarwal, M., Averett, M., Baumer, B.S., Bray, A., Bressoud, T.C., Bryant, L., Cheng, L.Z., Francis, A., Gould, R., Kim, A.Y., Kretchmar, M., Lu, Q., Moskol, A., Nolan, D., Roberto, R., Raleigh, S., Sethi, R.J., Sondjaja, M., Tiruvilumala, N., Uhlig, P.X., Washington, T.M., Wesley, C.L., White, D., Ye, P.: Curriculum guidelines for undergraduate programs in data science. *Ann. Rev. Stat. Appl.* **4**, 15–30 (2017)
22. Indiana University – Purdue University Indianapolis, School of Informatics and Computing Departments. <https://soic.iupui.edu/departments/>. Accessed 2 Sept 2017
23. Indiana University – Purdue University Indianapolis, School of Informatics and Computing Mission. <https://soic.iupui.edu/about/mission/>. Accessed 2 Sept 2017
24. Dataists: A Taxonomy of Data Science. <http://www.dataists.com/tag/osemn/>. Accessed 12 Sept 2017
25. DataONE: The Data Life Cycle. <https://www.dataone.org/data-life-cycle>. Accessed 12 Sept 2017
26. Muro, M.: Tech jobs are spreading into ‘Flyover Country’. *Wall Str. J.* (2016). <https://blogs.wsj.com/experts/2016/09/18/tech-jobs-are-spreading-into-flyover-country/>
27. Muro, M., Liu S.: Tech in metros: the strong are getting stronger. Brookings, 8 March 2017. <https://www.brookings.edu/blog/the-avenue/2017/03/08/tech-in-metros-the-strong-are-getting-stronger/>
28. Indiana Economic Development Commission: Governor Pence announces plan to invest \$1 billion in Hoosier innovation and entrepreneurship, 14 July 2016. [http://www.in.gov/activecalendar/EventList.aspx?fromdate=1/1/2016&todate=12/31/2016&display=Month&type=public&eventidn=249728&view=EventDetails&information\\_id=247868&print=print](http://www.in.gov/activecalendar/EventList.aspx?fromdate=1/1/2016&todate=12/31/2016&display=Month&type=public&eventidn=249728&view=EventDetails&information_id=247868&print=print)
29. Central Indiana Corporate Partnership: TechPoint workforce report: demand for computer-related jobs grows in Central Indiana as tech sector gains momentum, 14 March 2014. <https://www.cicpindiana.com/techpoint-workforce-report-demand-computer-related-jobs-grows-central-indiana-tech-sector-gains-momentum/>
30. Indiana University – Purdue University Indianapolis, School of Informatics and Computing, Data Studies Minor. <https://soic.iupui.edu/undergraduate/degrees/data-studies-minor/>. Accessed 2 Sept 2017
31. Indiana University – Purdue University Indianapolis, School of Informatics and Computing, Applied Data Science Minor. <https://soic.iupui.edu/undergraduate/degrees/data-science-minor/>. Accessed 2 Dec 2017