

## Introduction

**Judy Ruffenberg**, Director, Scholars and Scholarship, Association of Research Libraries

Data analytics and data science programs in research universities in Canada and the United States have grown dramatically over the past decade. An intrinsically interdisciplinary endeavor—drawing from statistics, computer science, engineering, and more—data science education has followed multiple trajectories. Projections of labor market demand for data manipulation skills, growing acceptance of open science and data sharing, the reproducibility “crisis,” and the promise of unlocking new scientific discoveries through analyzing massive amounts of data have all fueled enthusiasm for data science education. Funding models for new centers and institutes have multiplied, too—from large, personal, philanthropic donations (University of Virginia), to federal investment (Pan-Canadian AI Strategy and the Big Data to Knowledge, or BD2K, initiative), to private foundations (Gordon and Betty Moore and Alfred P. Sloan) that have invested in creating data science “environments” in three large US research universities.

This issue of *Research Library Issues* looks at the critical role and participation of libraries and librarians in supporting the data science revolution. **Catherine Murray-Rust, with Ameet Doshi, Jay Forrest, Ximin Mi, and Alison Valk**, demonstrates a pathway to data science through teaching core data-literacy skills to students. This is part of a broader library strategy at Georgia Tech to identify service gaps across the university that align with the library’s mission, and work with faculty and students to fill those gaps. But the library’s data-literacy program is doing much more than that. Through their offerings, including education on basic statistical packages and web scraping tools, the library is providing an informal opportunity for students in non-STEM fields to gain proficiency with these methods for their intellectual, professional, or vocational pursuits.

This is not surprising. “The library is the most interdisciplinary place on campus,” offered New York University (NYU) physics professor David W. Hogg, in his reflections on the Moore-Sloan Data Science Environment (MSDSE) at his university and its strong ties to the library. **Jennifer Muilenburg and Judy Ruttenberg** spoke with Hogg and other key personnel at the three MSDSE institutions—NYU, UC Berkeley, and the University of Washington—to highlight the profound growth and transformation of data science education that the MSDSE funding catalyzed. The article charts the instantiation of data science education at the three institutions, their aligned but distinct areas of focus, and key lessons for library leaders and other campus administrators in this arena.

Finally, **Erica Johns, Susan McCarthy, and Cynthia Parr** contribute two companion pieces on the highly collaborative growth of data science services at the US National Agricultural Library (NAL). Working with the University of Maryland iSchool and other land-grant university members of the Agriculture Network Information Collaborative (AgNIC), NAL identified scalable strategies to partner with the Agricultural Research Service (ARS) and the university research community to advance data-driven discovery in agriculture.

Data science is new enough as an endeavor that its relationship to its core disciplinary antecedents, and its infiltration of disciplines across the curriculum, is a dynamic and unfolding

“Research libraries...are prepared for this future as we draw upon our long-standing contributions: creating the conditions for new knowledge discovery, teaching students how to discern validity, and partnering with the research community to prepare and preserve data for science.”

story in higher education. Research libraries, however, are prepared for this future as we draw upon our long-standing contributions: creating the conditions for new knowledge discovery, teaching students how to discern validity, and partnering with the research community to prepare and preserve data for science.

© 2019 Judy Ruttenberg



This article is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

**To cite this article:** Judy Ruttenberg. “Introduction.” *Research Library Issues*, no. 298 (2019): 3–5. <https://doi.org/10.29242/rli.298.1>.