2020

Emerging Technologies for Research and Learning:Interviews with Experts

Scout Calvert

Mary Lee Kennedy

Clifford Lynch

John O'Brien

Emerging Technologies for Research and Learning: Interviews with Experts

By Scout Calvert and Mary Lee Kennedy

Edited by Clifford Lynch and John O'Brien

March 23, 2020

Association of Research Libraries
cni Coalition for Networked Information
EDUCAUSE

Table of Contents

Introduction

This report is a companion to *Mapping the Current Landscape of Research Library Engagement with Emerging Technologies in Research and Learning*,[1] which was commissioned by the Association of Research Libraries, the Coalition for Networked Information, and EDUCAUSE to survey research libraries' adoption and use of emerging technologies[2] to collaborate for the benefit of research and learning. The present report complements the landscape review by summarizing and synthesizing a series of interviews conducted in late 2019 and early 2020 with experts in research and academic technologies from several allied sectors: associations, businesses, academia, and the nonprofit research sector (see Appendix A). A list of emerging technologies and applications derived and expanded from EDUCAUSE research was a prompt for interviewees (see Appendix B). Interviewees were asked to address strategically critical advances in research and learning given emerging technologies, as well as potential opportunities and challenges for research libraries. The interviews were wide-ranging, surfacing insights about emerging technologies and their current and near-term applications in research and higher education environments.

Overarching Themes

Four broad themes emerged from the interviews:

1. It is indisputable that "data" has become a pervasive concern, paralleled by data-intensive science and scholarship. Already the basis of scholarly claims, data has long since escaped lab notebooks, ledgers, and journals, even if some of it lives an analog life on the way to a digital existence. Techniques for machines to use and generate data with less and less human intervention or interference continue to evolve, now more rapidly than in the past.
2. The concept of "smart machines," as defined by Shoshanna Zuboff in 1988, prompts a reminder that:

> 1) technology is not neutral, but embodies intrinsic characteristics that enable new human experiences and foreclose others, 2) within these new “horizons of the possible” individuals and groups construct meaning and make choices, further shaping the situation, and 3) the interplay of intrinsic qualities and human choices is further shaped by social, political, and economic interests that inscribe the situation with their own intended and unintended opportunities and limitations.[3]

We can expect machines to get better at being smart—and we can help shape their direction and adoption.

3. Our experts anticipate that research libraries will continue to function as partners in the enterprises of research and learning. They identify research library leadership opportunities in several areas: providing access to networks of data, increasing the availability of collections of data and related products and services, understanding the design and the implications of the emerging technologies so as to consult on their use, being the authoritative steward of data, and convening—and in some instances leading—discussions on the ethics of emerging technology.
4. To be successful, research libraries will need to negotiate both scale (applicability across disciplines and institutional boundaries) and context (applicability at the local level) in order to engage and lead. As perceived knowledge stewards, research libraries must work at scale to ensure sustainability of knowledge over time, even if it means redefining their economic and organizational models.

Primary Findings

Artificial Intelligence[4] and Data-Intensive Science and Society

Two suites of emerging technologies and practices were consistently identified by interviewees as strategically critical for research and learning: data and its management technologies, and artificial intelligence (AI) applications powered by data. Together, they compose the core of emerging technology applications for research and learning discussed in this report.

In the context of research collaborations, including emerging cross-disciplinary research, interviewees were most energetic about machine learning,[5] with its associated technologies and practices. Machine learning requires access to data and supports that scaffold access throughout the research life cycle, including, at times, cloud storage, computation and curation in place, data integrity checks, and high performance computing.

Factors That Influence Adoption of Emerging Technologies

Interviewees identified several factors that influence the adoption of emerging technologies in research and learning. They advised strategic exploration and adoption of new technologies, informed by the explicit values and objectives of the research and learning endeavor, and in consideration of the degree of alignment with the organization's mission.

Approach Strategy: Top Down or Bottom Up, the Focus Is Scale

There is an explicit tension in strategic approaches to adoption. For some experts, the approach that best works is to demonstrate that something works in practice and then to drive adoption from the bottom up. One interviewee stated that "every time we can do science and science infrastructure bottom up rather than top down, we save somewhere between five and ten years in the acceptance process. That's why I continually look for, what can a data repository or FAIR [Findable, Accessible, Interoperable, Reusable] data repositories do, so that we can claim it as an ancillary technology to demonstrate something that has been done." For other interviewees, adoption can also be influenced by top-down developments, for example, by federal funding policies and standards requirements.

Nevertheless, all interviewees agreed that successful adoption of emerging technologies is dependent on strong partnerships among those who have access to the required resources (for example, talent, finances, data). Building coalitions and proving economies of scale are essential.

Policy and Data Governance

Emerging technologies carry policy and regulatory implications. Several national and international initiatives, particularly as they relate to ethics, are already in place.[6] Among the experts we consulted, there was no consensus about what to prioritize regarding the future direction of related policy and regulation, though some expect to see further regulation of the internet and, particularly, personal information sharing. Others felt that university data governance can bring clarity as institutions grapple with data ownership and personal and research data security issues.

Privacy and Ethics

A major concern about the adoption and development of emerging technologies identified by several respondents was the appropriate handling of data privacy in research and learning. Several key emerging technology applications are tied up with privacy, including learning analytics and adaptive learning. Questions about privacy, informed consent, and ethics are unsettled as new methods for collecting and analyzing data emerge. Interviewees described a polarized privacy discussion, with advocates of absolute privacy on one side and those who believe potential benefits outweigh any risks on the other, and little nuanced discussion in between. This concern is directly related to the needed clarity on policy and governance noted above.

While student success initiatives hold promise for ensuring more students complete their college degrees, interviewees noted that conversations about the ethics of collecting and using data for this purpose haven't kept pace with emerging technologies. New kinds of tracking, from card swipe to internet-enabled scooters, underpin these data sets. As with predictive learning analytics, decision-making with student data has gained traction faster than clarity about the implications of collecting and using it. As one respondent noted, “There are certain kinds of protections that are in place through FERPA and so forth, but there are other kinds that simply are not what we’re really thinking about.” Students have not had a chance to consent to research, and the benefits of interventions based on the data have not been fully studied.

Technological Advancement

There is no single path for emerging technologies to advance, pause, develop, or recede. The interviewees reminded us that sometimes adoption is a “game of adjacency.” In other words, a technology may see funding and development postponed until a related technology advances, as in the case of high speed networks that are required to move data in the service of high performance computing centers. A new technology may also be adopted in a closely related field after another discipline has demonstrated its usefulness. A good example of this is the way machine learning has become so prominent, fueled by the availability of very large data sets and enormous computational resources. Locally, the implementation of new technologies or practices can be disruptive to productivity while the organization’s workforce adjusts workflows and becomes accustomed to new procedures. And, as with most changes, fear can stall adoption efforts.

Expertise and Labor

Given the emphasis in science on AI and data science, it is no surprise that there is an increased demand for expertise. Studies highlight the national, even global, focus on increasing expertise in artificial intelligence and its dependence on data, and data science. There is continued speculation about the AI workforce composition and impact; however, the demand for AI skills remains high. Universities have increased their data science degree programs, as well as opened new campus-wide data science and research data institutes.[7] In 2019 there was a significant shortage of data scientists in the United States that persists into 2020.[8]

The demand for data science skills has driven the establishment of data science courses available through a variety of instructional channels, from learning platforms like Coursera to certificate programs and master's degree programs at bricks and mortar institutions, with multiple disciplinary homes, including iSchools and computer science and statistics departments. One interviewee cited over 200 data science degree programs in the United States, as well as the emerging campus-wide data science centers such as the Berkeley Institute for Data Science[9] and the recent Academic Data Science Alliance[10] resulting from the Moore-Sloan Data Science Environments. For higher education, there is an opportunity to strengthen relationships with students and alumni for lifelong learning, offering microcredentials for continuous learning of specific competencies and advancing career goals.

Changing Demographics in Higher Education

Significant demographic changes in higher education are underway and are expected to continue over the next five years. These changes include enrollment declines for some institutions, a shift in student population across the United States to the west and southwest, and greater diversification of the student population.[11] Some of the experts noted that emerging technologies have the potential to improve or expand online education and support student success, for example, through AI-enabled learning and learning analytics.

Emerging Technologies in Learning and Research

Data intensive emerging technologies manifest in learning and research in several ways, offering routes to impact student success and transform the research process. In learning, emerging technologies may enable new strategies for addressing student success, with the full impact on learning yet to be understood. In research, emerging technologies are already impacting scientific methodologies, workflows, and consequently the nature of research, with more disruption, and opportunity, anticipated. As the nature of research changes, so do the practices surrounding its dissemination, including data and code sharing, with implications for reproducibility, new forms of peer review, and other issues taking on new importance.

Learning

In light of demographic shifts, we can expect student bodies to continue to diversify, not just in terms of gender and racial background but also in terms of other factors like first-generation college attendance. As student bodies become less homogeneous, there is an opportunity to develop emerging technologies to meet their specific learning needs.

Online learning continues to trend with these expanding technological capacities, as some campuses implement online-only degree programs. One interviewee predicts that "the most interesting applications in the five- to ten-year range are not going to be the on-campus versions [of learning], but it's going to be the distance learning versions of them."

Emerging technologies have the potential not only to provide educational opportunities asynchronously and at a geographic remove, but also to improve their efficacy and learning outcomes, both online and in physical classrooms. AI has the potential to enhance instruction in the classroom as well as online.

Research

According to the interviewees, scholars in every discipline can and will expect access to more data, whether purchased, gathered through local generation, or shared across networks, as well as new modes of aggregation and data-intensive research. Researchers expect to develop new methods to have more ways to produce and use data, as the "data deluge" reaches less data-intensive disciplines.

Artificial intelligence applications for automating research tasks are being developed in multiple quarters. AI and machine learning have the potential to automate discovery and assessment of scholarly literature,[12] update and implement research data management plans (machine actionable DMPs), and streamline or control other aspects of scientific workflows. The quantity of existing literature far exceeds the ability of human experts to synthesize and engage with it. Research in this area seeks to identify ways in which emerging technologies can accelerate scientific discovery by adopting artificial intelligence in research workflows, including the adoption of emerging technologies to automate the analysis of research papers and data.[13] One report suggests that it is unclear whether or not "automation of research will free up and empower researchers to be more creative and productive, or whether it will instead replace them entirely."[14]

There are many examples of the use of artificial intelligence technologies in scientific discovery, and data-centric knowledge products are becoming more common in academia. These developments highlight the importance of expert curation, which ensures access and trust in the provenance and representation of scholarly outputs—all key roles of research libraries that are further discussed below.

The experts we interviewed noted that data sharing and data citation are more frequent; data sets and data repositories are more visible to scientists and are common scholarly outputs. They noted that scholars share real-time data dashboards in their research and also use systems that collect data in real time and apply data stream algorithms and applications.

These findings framed the consequences for research libraries discussed with the experts.

Consequences for Research Libraries

As shown in the companion landscape report, research libraries are well on the path of data services delivery. Inter-institutional structures, such as the Data Curation Network, are in place, and some have relationships with data repositories such as Zenodo and Dryad, or provide a local data repository. As research and learning change with the adoption of emerging technologies, research libraries have an opportunity to collectively and individually define how these technologies advance their contribution to the academic mission—in skills development, or in providing related research services, or some combination of the two. The interviewees identified opportunities, and a sense of urgency, for research libraries to boldly shift their focus to data, expand their role in the research enterprise, scale and cooperate locally, provide access to AI-ready local and networked information, and partner in teaching and learning, all while leading with library values. Research libraries can build credibility as collaborators by being both users of emerging technologies and trusted advisors. However research libraries define the near future, it will require understanding the trade-offs in the context of their parent institutions' finances and partnerships, which may elicit new economic models that balance service delivery with financial sustainability.

Make the Strategic Shift to Data

As one interviewee put it, research libraries should "stop doing what you are already doing. We now have web documents. That's a done deal. People understand that. Your new challenge is to gain the expertise and the ability to do for science data what you've done for science text.... Bear in mind that a number of your new customers, readers of your material, will be computers, not humans." This exhortation speaks to at least two significant areas of opportunity for research libraries: in data stewardship and FAIR principles,[15] and in knowledge representation, curation, and structuring.

Recognizing that the support of provosts and other key decision makers is essential, interviewees speculated about whether research library leadership is ready to make the changes and whether or not they have the policies and decision-maker agreement needed to support them.

Per the interviewees, success will mean that research libraries:

- Steward data through the full life cycle of research, authoritatively and with accountability
- Consult on the analysis of data and rights regimes, such as licensing
- Reduce administrative burden for researchers to allow more time on research
- Work to reduce the cost of data curation and improve capabilities
- Collect and connect data sets and make it possible to discover and access them locally and globally
- Co-create data lakes and provide meaningful access to them
- Develop and maintain data-type registries and persistent identifiers for research data
- Provision data archiving tools
- Contribute new knowledge by studying the creation and use of new forms of data-based research and learning

Research libraries have a solid foundation on which to partner in the adoption of emerging technologies to advance learning and research, to conduct research using their own and other collections, and to scale up for collaborations that engage emerging technologies. But some upskilling will be required to meet partners confidently. Research libraries will benefit from in-house abilities to use the data they collect and generate. Expertise development will be required for research libraries that want to engage in this space.

Be Strategic Partners in the Research Enterprise

Interviewees stressed that these are still “early days” for data-intensive research, including related emerging technologies. Researchers are “trying to figure out how to use data to open up fields so they need to understand questions such as: ‘What’s the licensing regime? What’s the privacy regime? What’s the ethical use regime like?’ All those things have to be created as infrastructure to enable this to be useful. Libraries have played a big role in other transitions; it may be that libraries can play a role now."

As well as identifying new ways of delivering on their mission, research libraries will want to think strategically about how they can complement and substitute for current services. Our interviewees identified many opportunities for research libraries as partners in the research enterprise. Specific opportunities emerging from the interviews are outlined below.

Lead with Library Values

There is a significant need for the research and learning community to understand the realistic benefits and uncertainties of emerging technology applications. Experts we spoke to believe that libraries have the experience and capacity to convene informed conversations on data privacy, ethics of emerging technologies, and data rights retention. Crucially, research libraries were viewed as places (virtual and physical) where nuanced discussions about the nature of privacy and potential risks and benefits to users could take place. Engaging in these discussions, rather than avoiding them, was viewed as a necessary step for libraries to move forward with developing tools and platforms that promote student success, accelerate research, and enhance library services.

Research libraries are in an ideal role to act as ethics partners, particularly given their data management skills and knowledge of information ethics. They can work with researchers throughout the research process, including in relation to the use of third-party and tracking data. Furthermore, as ethical choices arise with further advancements in the use of emerging technologies, they can inform the ethical choices to be made. As one respondent put it, "I think librarians sort of bring a sensibility about data ethics that probably other folks don't naturally have."

Collaborate and Deliver at Scale

Research libraries cannot expect universal, off-the-shelf emerging technology applications; rather, participating in the development and application of emerging technologies will require libraries to develop expertise. Adoption and implementation will be tied into local contexts, from the specific needs of a variety of students and faculty, to the diverse modes of collaboration required by researchers.

Emerging technologies will give rise to opportunities to engage with other campus and research institution units. Research libraries have new opportunities to partner with campus units to acquire, manage, curate, and provision data and support its use. For example, libraries were cited as potential coordinators of service bureaus at their institutions that would bring together cross-campus research and learning services and include library expertise in metadata, research data management, and repositories. These service bureaus would include individuals with expertise in statistics, geographical information systems,, data cleaning, analysis, curation, coding, and data science tools and instruction. As one respondent envisioned, "If the data sets, and the access to the data, and the computing power needed to process it become tomorrow's library—I mean, then they're at the cutting edge."

This level of engagement is as true on campus as it is across geographies and disciplines. Such collaboration could develop into a deeper cooperative model for data repositories and research support services, including related standards setting and structured data, and engaging in interoperability and data science discussions. As one interviewee noted, "The name of the game is scale. What will the great library of 2030 look like?"

Lead in Research Information Access and Management

While many research libraries may already be active on both sides of collecting and disseminating data, they can expect to see both supply and demand increase for local and global use by scholars and students.

The maturity and co-evolution of storage technologies, networks, and computation make it possible to vividly reimagine the work of research libraries to develop “monumental new services” for research and learning. We can expect machines to get “smarter” and per our interviews, libraries are an “incredibly fertile area for that to happen because they have voluminous amounts of data.” Much of this data remains locked up in paper, software, or digital image files. Emerging technologies have the potential to release this data for use by machines in the service of research and learning. Collaborative initiatives can scale to support cross-disciplinary and international research projects.

Research libraries can play a critical role in shaping and influencing the design of data repositories so that the individual systems behave and appear as part of the larger community of data repositories in spite of the disparity in content, location, and governance. A series of design choices and practices will lead to meaningful interoperability. These include the following:

1. A persistent unique identifier resolution system. See, for example, the Digital Object Identifier (DOI) system used in the publishing industry.
2. Repositories that expose at least one interface (possibly among several agreed standards) that is based on semantic-free persistent unique identifiers. This provides a means of withstanding technological change. See, for example, the experimental academic use of a Digital Object Interface Protocol (DOIP).
3. Repositories that intrinsically support the execution of vetted operations on data contained within, and enable users to execute privileged operations, with adequate concern for individual privacy, rather than only enabling a request/response interaction pattern.

Research libraries can both create AI-ready collections from their local materials and collect new forms of data that are machine readable. The potential opportunities cited by interviewees include applications for processing digitized materials, machine-readable collections, deep content discovery, disability access, and linked data and cataloging. They highlighted the opportunity for research libraries to expand access to previously less accessible information, including the recognition of handwriting, faces, speech, images, and figures.

Research libraries can contribute by leading institutional pilots or collaborations, for example on FAIR data principles,[16] data repositories, or the provision of AI-ready collections to drive broader policy development. They could provide resources or experts who understand semantics and can advise scholars on AI as they look to ask questions of books, video, research data, and other holdings.

Partner in Teaching and Learning

Interviewees identified several opportunities in teaching and learning, inviting research libraries that may be waiting on the sidelines to move into the game. Discussion frequently reflected a view of the research library as an open, multidisciplinary, and familiar place (virtual or physical) where students and faculty generally expect to find assistance with a variety of technologies used in teaching, learning, and research. As more digitally enabled methods and tools become cross- or multidisciplinary, libraries have the opportunity to include these in their instructional repertoires, be it through information, data, and algorithmic literacy instruction sessions, or by establishing or partnering in the establishment of data science centers, building on the success of digital scholarship labs and maker spaces.

To advance their learning mission, research libraries can not only teach about emerging technologies, but also use them ethically and robustly in service of that mission. Specifically, research libraries have a role:

- In assisting transitions for scholars and students as they use more data, to facilitate interpretation, use of tools, and understanding of privacy, security, and ethics
- As a forum for discussing and demonstrating the broad range of ethical uses of emerging technologies
- In learning how to interpret data, use the tools, know how to

navigate complex digital spaces, and know what the limits are in terms of ethics, technical design, scientific rigor

The experts we interviewed believe that research libraries have a crucial role in teaching new technologies and data science methods, because libraries know "how to develop services that accelerate people's ability to either do this work or understand how this work is done or as a student to learn how to do data science."

Establish New Research Library Service Models

On the near horizon, our findings suggest that technology trends will lead to the following potential service models that merit considerable attention by research libraries:

- Enabling not just self-service operations that improve convenience and quality, but also machine delegation of tasks.
- Machine to machine-facing services, including access to computation-enabled content, computational programming, and machine outputs
- Expanded service capacity to both supply large sets of data and join in the demand for data

Concurrently with filling these potential roles, research libraries can explore and develop an economic model for service delivery that both provides for and benefits from these new technological configurations—locally and at scale. The experts we interviewed also believe that research libraries can add to their value by conducting research and analysis of the nature of research and learning under emerging technology conditions. As one interviewee said, the key goal is developing a service model that delivers the expected quality of services "while keeping the money in the library funded on the things that really need doing."

Conclusion

As data becomes more pervasive across society, science, and scholarship, alongside the technologies and practices that collect, utilize, and create it, research libraries have a window of opportunity to use emerging technologies to advance the mission of research and learning. As scholars and students engage with emerging technologies, research libraries can apply new configurations of technologies as active players in student success initiatives, and in service to data-intensive research. The experts we interviewed saw very significant roles for research libraries, and opportunities as collaborative partners.

Appendix A

Research and learning technology experts from the following institutions participated in interviews in late 2019 to early 2020:

Board on Research Data and Information,
National Academies of Sciences, Engineering, and Medicine
George Strawn, Director

Computing Science and Telecommunications Board
Jon Eisenberg, Director

Corporation for National Research Initiatives
Robert E. Kahn, Chairman, CEO, and President

Georgia State University
Phil Ventimiglia, Chief Innovation Officer and independent technology futurist

Indiana University
Bradley C. Wheeler, Vice President for Communications and Marketing, Vice President for Information Technology and CIO, and James H. Rudy Professor of Information Systems

Internet2

ITHAKA
Kevin Guthrie, President

Michigan State University
Kathleen Fitzpatrick, Director of Digital Humanities and Professor of English

National Academies of Sciences, Engineering, and Medicine

Pew Research Center
Lee Rainie, Director of Internet and Technology Research

Rutgers, The State University of New Jersey
Michele Norin, Senior Vice President and Chief Information Officer

UC Berkeley

Jenn Stringer, Assistant Vice Chancellor for IT and Deputy Chief Information Officer

UCLA

UC San Diego

Vince Kellen, Chief Information Officer

University of Virginia

Philip E. Bourne, Founding Dean of School of Data Science, Stephenson Professor of Data Science, and Professor of Biomedical Engineering

Appendix B

ARL, CNI and EDUCAUSE Partnership on the Role of Research Libraries in a World Shaped by New Technologies

List of Emerging Technologies in Research and Learning

September 30, 2019

Artificial Intelligence

- Augmented analytics and data discovery
- The rise of very large data resources, some of which are much too big to be transferred over the network in any kind of reasonable time
- Machine learning
    - ► For machine learning (ML), think both about the implications of people wanting to run machine learning models on data (including very large data) and technologies enabled by ML such as face recognition, handwriting recognition, machine translation, authorship and plagiarism detection, etc. Also think about the need for training and reference data (e.g., databases of faces & names)
- Predictive analytics
- Text/content mining and analytics
- Automatic text generation (automatically writing news articles, scientific papers, etc.)
- Visual data discovery (visualization)
- AI for fielding transactional questions
- AI for wayfinding
- Deep fakes and re-animated historic figures

Research and Scholarship

- Author identifiers (ORCID, Researcher ID, Scopus). More generally, broad use of machine-processable IDs for research objects and entities of all kinds

- Structured machine-parsable data: knowledge graphs, linked open data, computer-parsable factual biographies, etc.
- High performance computing, including specialized computational resources (GPUs, FGPA-based machines, quantum machines, etc.)
- Data repositories: institutional, disciplinary, governmental. Rise of scientific information platforms (that include services, not just static data at rest). Scientific workflow systems
- Very high performance networks for science: Pacific Research Platform and National Research Platform developments. High performance and easy-to-build scalable simulations. Research data sharing across institutions and internationally
- New models of co-locating very large data and computation; sending computation to the data rather than importing data to a local computational environment
- Virtual, augmented, mixed reality
- Natural language processing
- Containerization/Kubernetes and emulation
- High precision geolocation services (related to augmented reality, among other things)
- Smart campus and smart cities; ubiquitous sensors

Security and Privacy

- Identity as a Service (IDaaS), including identity via social media ("Sign on with Facebook")
- Authentication; potential end of passwords
- Large scale identify federations, nationally and internationally
- Machine learning and autonomous security-based models
- Differential privacy

Social/Personal/Communication

- Technology to make systems and content accessible
- Speech recognition
- Social bots (e.g., chatbots)
- New developments in collaboration platforms (Slack)

Teaching and Learning

- Adaptive learning
- AI-driven instruction
- Technology enhanced advising
- Electronic learning platforms
- Learning analytics
- Games and gamification
- Digital microcredentials
- Incorporation of mobile devices in teaching and learning
- Open education resources (e.g., textbooks, data)
- Robotics
- Data science, computational intensive instruction
- Uses of the Internet of Things for teaching and learning

Others

- Video analysis (e.g., Labelling and segmentation, automatic closed captioning)
- Robotic process automation
- Drones
- Internet of Things
- Blockchain and distributed ledger systems
- Commercial clouds (AWS, Google, IBM, Microsoft Azure). Multi-cloud systems; implications of lock-in, ingress, and egress charges
- G5 wireless technology; edge processing for low-latency services

Endnotes

1. Sarah Lippincott, *Mapping the Current Landscape of Research Library Engagement with Emerging Technologies in Research and Learning*, ed. Mary Lee Kennedy, Clifford Lynch, and Scout Calvert (Association of Research Libraries, Born-Digital, Coalition for Networked Information, and EDUCAUSE, forthcoming); *Executive Summary* published March 2020, https://doi.org/10.29242/report.emergingtech2020.landscapesummary.

2. The definition of emerging technologies developed by Rotolo, Hicks, and Martin "identifies five attributes that feature in the emergence of novel technologies. These are: (i) radical novelty, (ii) relatively fast growth, (iii) coherence, (iv) prominent impact, and (v) uncertainty and ambiguity."—Daniele Rotolo, Diana Hicks, and Ben R. Martin, "What Is an Emerging Technology?," preprint, submitted February 13, 2015, last revised January 4, 2016, https://arxiv.org/abs/1503.00673.

3. "In the Age of the Smart Machine," Shoshana Zuboff's website, accessed February 25, 2020, https://shoshanazuboff.com/book/books/in-the-age-of-the-smart-machine/.

4. For the purposes of this paper we use the following definition of AI from the Association for the Advancement of Artificial Intelligence (AAAI): "the scientific understanding of the mechanisms underlying thought and intelligent behavior and their embodiment in machines."—AAAI, "Information about AI from the News, Publications, and Conferences," *AITopics*, accessed February 25, 2020, https://aitopics.org/search. One can think of AI as a large collection of interrelated technologies; in the last few years, AI has often and misleadingly been treated as a synonym for machine learning.

5. Here, we use a straightforward definition: "Machine learning is the science (and art) of programming computers so they can learn from data."—Aurélien Géron, *Hands-On Machine Learning with*

Scikit-Learn and TensorFlow (Sebastopol, CA: O'Reilly Media, 2017). A subfield of AI, Pal and Gulli explain that machine learning "focuses on teaching computers how to learn without the need to be programmed for specific tasks."—Antonio Gulli and Sujit Pal, *Deep Learning with Keras* (Birmingham, UK: Packt Publishing, 2017).

6. Mary Lee Kennedy, "What Do Artificial Intelligence (AI) and Ethics of AI Mean in the Context of Research Libraries?," *Research Library Issues*, no. 299 (2019): 6, https://doi.org/10.29242/rli.299.1.

7. For a list of campus and university data science programs see "College & University Data Science Degrees," DataScience. Community, accessed February 25, 2020, http://datascience.community/colleges.

8. Jen DuBois, "Plan to Hire Data Scientists in 2020? Here's What You Need to Know," QuantHub, December 19, 2019, https://quanthub.com/hire-data-scientist/.

9. Berkeley Institute for Data Science, accessed February 25, 2020, https://bids.berkeley.edu/.

10. Academic Data Science Alliance, accessed February 25, 2020, http://www.academicdatascience.org/.

11. Nathan D. Grawe, *Demographics and the Demand for Higher Education* (Baltimore, MD: Johns Hopkins University Press, 2018), https://jhupbooks.press.jhu.edu/title/demographics-and-demand-higher-education.

12. Wolfgang E. Kerzendorf, "Knowledge Discovery through Text-based Similarity Searches for Astronomy Literature," *Journal of Astrophysics and Astronomy* 40, no. 3 (2019): 23, https://doi.org/10.1007/s12036-019-9590-5.

13. The National Academy of Sciences is currently focused on this very topic. See "Public Workshop: Opportunities for Accelerating Scientific Discovery: Realizing the Potential of Advanced &

Automated Workflows, March 16–17, 2020," National Academies of Sciences, Engineering, and Medicine, accessed February 25, 2020, https://sites.nationalacademies.org/PGA/brdi/realizing-opportunities-for-advanced-and-automated-workflows-in-scientific-research/PGA_198528.

14. Elliot Jones, Nicolina Kalantery, and Ben Glover, *Research 4.0: Interim Report* (London: Demos, October 2019), https://demos.co.uk/wp-content/uploads/2019/09/Jisc-OCT-2019-digital-final.pdf.

15. Mark D. Wilkinson et. al, "Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardship," *Scientific Data* 3 (March 2016), https://doi.org/10.1038/sdata.2016.18.

16. "FAIR Principles," GO FAIR, accessed February 25, 2020, https://www.go-fair.org/fair-principles/.