ACRL PUBLICATIONS IN LIBRARIANSHIP NO. 78

The Rise of AI

Implications and Applications of Artificial Intelligence in Academic Libraries

edited by
Sandy Hervieux and
Amanda Wheatley

ACRL PUBLICATIONS IN LIBRARIANSHIP NO. 78

The Rise of AI

Implications and Applications of Artificial Intelligence in Academic Libraries

edited by
Sandy Hervieux and
Amanda Wheatley

Association of College and Research Libraries
A division of the American Library Association

Chicago, Illinois 2022

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992. ∞

Library of Congress Control Number: 2022933641

Printed in the United States of America.

26 25 24 23 22 5 4 3 2 1

Contents

Acknowledgments

The authors would like to acknowledge the support of the following institutions and individuals during this project: the McGill University Library and our colleagues at the Humanities and Social Sciences Library, Erin Nevius, and the Association of College and Research Libraries. The editors also wish to thank the chapter authors for all their hard work and their willingness to share their innovative ideas, projects, and initiatives. A special thank you to our families for their unwavering support and encouragement: Sylvie et Robert, Tammy, Anis, Melanie, and Laura.

Introduction

Artificial intelligence or AI, if you prefer, has been in the minds of researchers and dreamers alike for decades. It brings about connotations of science fiction and fantasy, but perhaps more importantly, of progress. After all, AI and progress go hand in hand. So where then does the academic library fit into the ever-growing expanse of artificial intelligence? The answer to this question was posed over forty years ago as librarians began to see the shift from automation to computer intelligence. Nearly half a century has been spent trying to prepare for a hypothesized takeover of librarian jobs. This fear would seem natural: librarians are information stewards, and some of the most brilliant technological advancements in the last hundred years have sought to open information access to an even broader public. The addition of AI to these advancements has made the information seeking process nearly unrecognizable from previous generations.

So, what exactly is artificial intelligence? The editors of this book provide a living definition that they believe stands true at this time; however, they recognize the fluidity of the field and the ability for this definition to change over time. Ask them again ten years after the publication of this book or even after a few months, if you prefer, and it is likely that the changing landscape of AI will have brought on new considerations to this definition. Nevertheless, for the purposes of this book, the editors define artificial intelligence as the development of machines to accomplish tasks and reproduce thought processes that are normally seen in humans; this simulation of intelligent behaviour is unique from other automation as it requires the computer to use human reasoning or thinking to perform tasks.

Along with AI, machine learning has quickly provided new solutions to information professionals. Machine learning (ML) is acknowledged as a subset of AI, though some scholars and scientists would argue it is its own field. The application of machine learning begins with its ability to learn from data and make decisions without the express intervention of a human. An ML algorithm can observe and detect patterns in data with the goal of being able to predict future decisions and outcomes from said data. ML learns as it goes and can adapt new stimuli into its decision-making process, much like that of a human. The advancement of ML technologies opens the door for endless possibilities for librarians to utilize these programs to classify, label, and organize information in an automated process. Some of these possibilities will be explored within the chapters of this book, providing innovative new ways to integrate machine learning and artificial intelligence into the work of librarians.

Much in the same way that Johannes Gutenberg's invention of the printing press revolutionized the public's access to information, artificial intelligence has done the same. Search engines and search algorithms are the printing press of the day. The monopolization of Google over information seeking and point-of-entry access has given millions the ability to seek out and discover information at rates heretofore unseen. The Google search engine has only been made more powerful by the improvements and inclusion of AI in its design. Without AI, a search engine is nothing more than a simple if/then statement. If the user searches for X, retrieve them Y. Of course, we all know that this is not how a search engine using AI functions. For most of the public, it is enough to understand that the code can interpret a semantic meaning for X, allowing the search engine to retrieve results for Y but also for Z. The proverbial AI *black box* allows users to grasp the concept that intelligent code exists within the box and that code can learn as it goes, similar to a human.

Librarians are uniquely positioned to rise to the challenge that AI presents to their field. Libraries and their like have existed for millennia; they progress with society, altering and adapting their services to meet the information needs of their communities. Academic libraries today have greatly expanded their digital offerings, not just to include electronic books or journal articles but also to support software application discovery and use. Some academic librarians might say they lack a foundational knowledge of AI or that they are ill-equipped to speak on the subject, and yet they have likely been interacting with AI through the different types of software applications they support. At the very least, they have encountered and mastered the art of the search algorithm. Librarians do not need to be proficient in the contents of the *black box* to provide AI support, as this book will show, they need only the desire to try. If that is not inspiration enough, consider library users struggling to locate a book through signage and maps. Would they be refused help just because a librarian didn't understand cartography? Librarians need not be computer scientists to partake in the conversation surrounding AI, they only need to be curious.

What follows in this book is a snapshot in time, a look at the academic librarians who have risen to the occasion and have begun to embrace AI in their work. The book is organized in three parts to establish AI initiatives in User Services, Collections and Discovery, as well as exploring the movement in Toward Future Applications. As users confront their own understanding of AI, some librarians have reached out to establish communities of discourse, such as The 99 AI Challenge, Keeping Up with Artificial Intelligence and workshops on machine translation. Others have created incubation spaces such as The Collaboratory and an AI Lab. Even further, some academic librarians have gone so far as to confront their users with robotics. Although some could argue that public libraries would be better suited to introducing their users to artificial intelligence, academic libraries are in a unique position where they can combine their information literacy initiatives with AI literacy. They can also foster interesting partnerships with different on-campus groups that can directly benefit their communities.

Behind the scenes, academic librarians have been integrating AI technology into their collections and using it to enhance discoverability. The use of AI to refine metadata for images, articles, and theses has played a large role in the improvement of university collections and institutional repositories. Collaborations with third-party groups have also enabled academic librarians to explore the use of AI applications in hand-text recognition through Transkribus and machine learning through IBM's Watson. While the nature of metadata and cataloging make them natural fits for AI, the importance of training librarians and staff members in this technology is paramount. AI can enhance the discoverability of collections and items, but the implications of its use should be understandable to its users.

Although practical applications of artificial intelligence are growing in academic libraries, much of the possibilities of this technology remain theoretical. The book ends with two chapters that explore the future possibilities of AI for academic libraries. As the discussion surrounding ethics, bias, and privacy in AI continues to grow, libraries will be called to make informed decisions and position themselves as leaders in this discourse. Another important implication for librarians will be how AI will impact information behaviour and how they must be aware of machine information behaviour and its main tenants.

In chapter 1, The 99 AI Challenge: Empowering a University Community through an Open Learning Pilot, authors Carey Toane, Lise Doucette, Paulina Rousseau, Michael Serafin, Michelle Spence, and Christina Kim explore the success of the first ever AI community workshop program at the University of Toronto.

In chapter 2, URI Libraries' AI Lab- Evolving to Meet the Needs of Students and Research Communities, authors Harrison Dekker, Angelica G. Ferria, and Indrani Mandal discuss the creation of the first AI Lab in an academic library in the United States of America.

In chapter 3, Artificial Intelligence, Machine Translation, and Academic Libraries: Improving Machine Translation Literacy on Campus, authors Lynne Bowker, Maria Kalsatos, Amy Ruskin, and Jairo Buitrago Ciro provide details on a project to support machine translation for international students through literacy instruction.

In chapter 4, Incubating AI: The Collaboratory at Ryerson University Library, authors Fangmin Wang, Aaron Tucker, and Jae Duk Seo discuss the creation of The Collaboratory, a multi-disciplinary research space at a Canadian university. The formation of the space, examples of scholarship produced from its collaborations, and services offered to the public are detailed.

In chapter 5, Separating Artificial Intelligence from Science Fiction: Creating an Academic Library Workshop Series on AI Literacy, authors (and editors of the book) Amanda Wheatley and Sandy Hervieux from McGill University in Canada describe the development of their three-part workshop series Keeping Up with Artificial Intelligence as a first step in instigating conversation around AI literacy within their institution.

In chapter 6, Do Students Dream of Electric Cats (or Dogs)?: Using Robotics for a Unique Exam Week Activity in the Library, authors Jonathan Scherger, Juliana Espinosa, Autumn Edwards, Chad Edwards, Bryan Abendschein, and Patricia Vander Meer discuss the success of using robotic animals as a replacement for therapy pet visits in outreach ventures in the library.

In chapter 7, Subjectivity and Discoverability: An Exploration with Images, authors Catherine Nicole Coleman, Claudia Engel, and Hilary Thorsen from Stanford University look at the results of using machine learning to classify images and improve the accessibility of collections.

In chapter 8, AI-Informed Approaches to Metadata Tagging for Improved Resource Discovery, authors Charlie Harper, Anne Kumer, Shelby Stuart, and Evan Meszaros explore the use of unsupervised machine learning to improve the discovery of electronic theses and dissertations.

In chapter 9, "We could program a 'bot' to do that!": Robotic Process Automation in Metadata Curation and Scholarship Discoverability, authors Anna Milholland and Mike Maddalena examine how robotic process automation can be used to establish metadata for an institutional repository at William & Mary Libraries.

In chapter 10, More Than Just Algorithms: A Machine Learning Club for Information Specialists, authors Mark Bell and Leontien Talboom from the National Archives in the United Kingdom discuss the formation of a Machine Learning Club within the archives and its use in bringing together professional and scholarly conversations on the use of ML in information sciences.

In chapter 11, The Role of the Library When Computers Can Read: Critically Adopting Handwritten Text Recognition (HTR) Technologies to Support Research, Melissa Terras evaluates the use of Transkribus software to support the digitization of handwritten text in libraries.

In chapter 12, Using IBM Watson for Discovery and Research Support: A Library-Industry Partnership, authors Aaron Trehub and Ali Krzton from Auburn University Library provide details on their use of IBM's Watson to apply machine learning to institutional research discovery.

In chapter 13, Ethical Implications of Implicit Bias in AI: Impact for Academic Libraries, authors Kim Paula Nayyer and Marcelo Rodriguez discuss the ethical considerations for libraries surrounding AI, with specific attention on the implicit bias of machine learning algorithms.

And in chapter 14, Machine Information Behaviour, author Michael Ridley delves into the implications of machine information behaviour as a launching point for further discovery and use surrounding machine learning in the information sciences.

The goal of this book is not prescriptive; it aims to introduce librarians to certain implications and applications of artificial intelligence, to start conversations, and to inspire. As the presence of artificial intelligence continues to grow in society, libraries will have to contend with the place they want to take with regard to this technology. The chapters in this book show the endless possibilities for librarians to engage with AI.

Part I
User Services

Chapter 1

The 99 AI Challenge:

Empowering a University Community through an Open Learning Pilot

Carey Toane, Lise Doucette, Paulina Rousseau, Michael Serafin, Michelle Spence, and Christina Kim

Introduction

While AI has entered the mainstream discourse in the past decade, teaching and research focusing on artificial intelligence are not new to the University of Toronto (U of T). As the home of the Vector Institute for Artificial Intelligence and the "godfather of deep learning," Geoffrey Hinton, U of T, and the city of Toronto have a rich history in the field of artificial intelligence (AI) and have become a hub for AI research.[1] However, many faculty, staff, and students are still in the dark about what AI is, how it works, and what implications it might have on their lives. While many departments at U of T offer AI courses and events, they are often aimed at those already engaged in this area of research or study who have a technical understanding of AI.

The library is in a unique position to facilitate equitable and open community conversation and learning in this area. Consisting of more than forty libraries spread over three campuses in the Greater Toronto Area, University of Toronto Libraries (UTL) is committed to supporting teaching, learning, and research through its world-class collections and services. Library support for AI research has recently grown to include business, engineering, entrepreneurship, law, and humanities as AI applications expand beyond the scope of computer science.

Born out of a library interest group and inspired by a similar experiment in Finland, the 99 AI Challenge was a year-long, community-based pilot to build capacity and awareness of AI among non-technical experts from a variety of backgrounds at U of T.[2] The libraries' commitment to teaching information literacy, to critical thinking, and to the evaluation of information, as well as the flexibility to offer cross-disciplinary programming, made it a natural fit. In addition to providing an opportunity for sustained engagement with users on a topic of key importance, the library also stands to benefit from a deeper understanding of the user community's position on this technology as the library adopts AI in its practices.

This chapter explores the experience of the administrators and participants in this pilot and evaluates the extent to which it met its goals. Additionally, the authors seek to understand its impact on the cohort's perceptions and attitudes toward AI as well as any recommendations for its function, adoption, and/or use.

Literature Review

AI traditionally falls within the realm of computer science education,[3] where there has been a historical disparity both in gender and race. The *2017/2018 Digest of Education Statistics* (US) indicates that men accounted for 80% of the total recipients of computer undergraduate degrees conferred in that year, with 55% of the male students identifying as white.[4] That year, 73% of students enrolled in Canadian mathematics, computer, or information science programs identified as men.[5] Singer writes that highly competitive programs across the US "disadvantage people who are already unrepresented in computer science—including women, African-Americans, Latinos and low-income, first-generation college students."[6] This disparity is perpetuated in the workforce, where women and minorities are underrepresented at major tech companies such as Apple, Facebook, Google, and Microsoft, as well as at smaller startups.[7] A recent U of T study examined the backgrounds of 585 founders of 335 Canadian tech companies and found that only 5.8% were female, and only 12–14% had a humanities or social sciences undergraduate degree, with engineering graduates dominating the field at almost 35% of founders.[8]

In a survey of AI educators and practitioners, Wollowski et al. found that only 16% of courses prioritized societal impact, ethical concerns, or philosophical issues in their goals or outcomes, although 41% did cover these topics at some point in the course.[9] Goldsmith and Burton use a case study on weaponized AI to demonstrate that ethical theory education is necessary for AI practitioners because it introduces new critical tools for decision making that the software may ultimately make on their behalf.[10] Goel makes an argument for AI education to be available to "all citizens so that they can make informed decisions about AI technologies without regard to hype about the wonders of AI wonders or unfounded fears of imagined threats,"[11] stating that the long-term success of the AI enterprise "will require the support of an informed citizenry."[12]

There appears to be limited literature connecting AI to librarianship. An environmental scan of twelve American and fifteen Canadian research library websites found no mention of AI in any strategic planning documents. Only 18.1% of the institutions offered any programming related to AI (which included workshops on coding and Arduino), and "startlingly few academic libraries have begun to engage in official projects or labs centered on artificial intelligence."[13] Wheatley and Hervieux also found that very few academic libraries offer AI programming or initiatives. They state that given that AI will have a great effect on the way that users search for information, libraries are well-positioned to lead AI instruction.[14] Arlitsch and Newell observe that libraries will become centres for continuing education in the area of AI, preparing both their staff and their communities for the coming changes.[15] A number of library associations have studied libraries' relationships to AI. A 2018 paper from the Canadian Federation of Library Associations encouraged the community to take the lead in advocacy and shaping discussions on AI.[16] The American Library Association states that AI may become another technological development that libraries help communities better understand.[17] In 2019, the Urban Libraries Council (ULC), a membership organization of North America's leading public library systems, expressed a vision where libraries help to serve their communities by advancing algorithmic literacy due to the potential impact of AI.[18]

Program Design

The AI Challenge launched in May 2019 upon receipt of a UTL Chief Librarian's Innovation Grant. The annual grant program identifies and provides funds for "innovative library projects that have the potential to be transformative for UTL and the communities that we serve."[19] An administrative team of six librarians and a faculty advisor led the program in outreach and communications, marketing and design, session facilitation, and data tracking. Funds were put toward promotional materials, catering, transportation, incentives, and speaker gifts.

The project consisted of two main phases: completion of an online course exploring the elementary knowledge of AI as well as an accompanying Slack discussion group (Phase 1: July–August 2019) and six in-person group conversations with U of T subject experts (Phase 2: September 2019–March 2020). In order to graduate from the challenge, participants were asked to complete the online course, attend four of the six conversation sessions either in-person or asynchronously by watching and responding to a video recording posted on the Slack channel, and complete two feedback surveys. Those who met all the milestones received a certificate and a chance to win one of five e-readers. Students could also receive an official co-curricular record from the university.

Recruitment ran for one month during May and June 2019. The team distributed printed and digital materials to an extended network of contacts, with the support of library and centralized university communications. From its inception, the challenge aimed to include those who may otherwise not have access to this kind of educational opportunity, in alignment with the university's statement of commitment to diversity.[20] Special efforts were made to identify potential student applicants through targeted email invitations to campus clubs and groups. Similarly, the admin team reached out to hospitals, schools, and government to recruit community members with a connection to U of T. All materials lead to a library webpage with an introduction, a FAQ, and an application form.

There was a high level of interest, with approximately 550 applications from current students at all levels, staff and faculty across disciplines, and community members. Applicants were screened for level of existing knowledge on the topic as well as to create an inclusive and diverse group. In the selection process, points were assigned according to a rubric. The admin team selected ninety-nine individuals to reflect approximately 0.1 percent of full-time enrolment: a cohort consisting of fifty-five students, fifteen faculty, fifteen staff, and fourteen community partners.

Phase 1 began in June 2019, when successful participants attended meet-and-greets on all three campuses where they received branded AI Challenge t-shirts. The admin team selected an open educational resource (OER) called Elements of AI,[21] created by the University of Helsinki for widespread implementation among non-technical participants in that country. Available in English, the OER has six modules encompassing definitions of AI and its related fields, philosophy, applications, functionality, and societal implications. The admin team guided the progress through the modules with weekly emails of encouragement as well as parallel peer-to-peer discussions over the Slack channel. Phase 1 wrapped up in August with the conclusion of the course, after which participants sent in their completion certificates and responded to the Phase 1 survey.

Phase 2 took place through the fall and winter and consisted of six events. U of T experts presented on the use of AI in immigration policy and practice (September 2019), the ethics of AI (October 2019), the potential for AI to redress inequalities in healthcare (November 2019), the interplay of art and performance and AI (January 2020), machine learning bias and the perpetuation of racism and discrimination (February 2020), and the realities of AI startups by a panel of women founders (March 2020).

These two-hour, in-person sessions were designed to facilitate active peer-to-peer learning rather than lecture-style instruction. The speakers gave fifteen- to twenty-minute presentations, after which the cohort was guided through a sixty-minute conversation facilitated by a librarian. Speakers provided three questions that the group responded to using a think-pair-share model. Speakers were encouraged to stay and join the conversation. The remaining thirty minutes of the sessions were dedicated to debriefing and informal networking. Those who were unable to attend a session in person were able to participate and receive attendance credit by watching a video of the presentation and responding to the conversation questions via the Slack channel.

Phase 2 concluded in March 2020 with the distribution of the second survey. An in-person graduation ceremony was postponed indefinitely due to COVID-19-related building closures. Alternatively, the graduates were announced via Slack, digital certificates were delivered via email, and draw prizes were mailed to the winners.

Assessment

Two online surveys were administered to participants: Survey 1 (abbreviated as S1) in August 2019 at the end of Phase 1 and Survey 2 (abbreviated as S2) in March 2020 at the end of Phase 2. All challenge participants were invited to participate in the activities in both phases and both online surveys/reflections, regardless of whether they met previous milestones.

Data collected included demographic information; reflective content such as learning assessment, change in perspective or opinion, future application of gained knowledge, and remaining questions to be explored; and program assessment information such as satisfaction with the content, timing, and design of the elements of the program, and willingness to recommend the program. The Phase 1 survey included twenty-four questions organized into three sections; the Phase 2 survey included thirty-four questions in five sections.*

Cohort Profile

Demographic information is based on respondents who completed the first survey. Figure 1.1 shows that just over half of respondents were students (57%) (S1 Q19, n=80), and just over half were between 18 and 34 years old (59%) (S1 Q24, n=79).

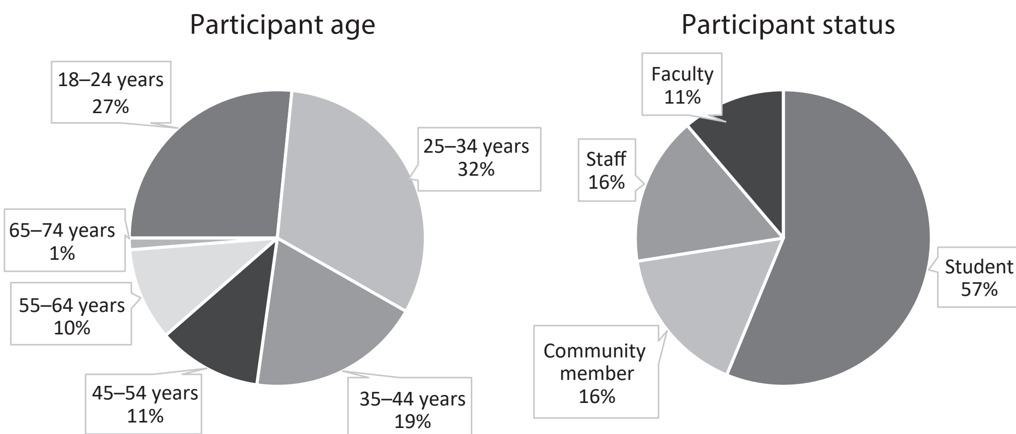


Figure 1.1

Demographics of Survey 1 respondents (age: S1 Q24 n=79; Status: S1 Q19, n=80).

Faculty members came from the Faculties of Arts and Science, Information, Management, Medicine, Music, and Social Work, and staff came from various academic and administrative units. Students came from the Faculties of Arts and Science, Applied Science and Engineering, Medicine, Music, Education, Information, Management, Law, and Public Health (S1 Q21, n=76). Survey respondents self-identified as belonging to one or more of the groups in table 1.1, reflecting the diversity of the cohort.

* The survey instruments and the program application form are available for download at http://hdl.handle.net/1807/101825.

Table 1.1

Demographics of Survey 1 respondents (identity: S1 Q23, n=79)

Group	Percentage
Female or non-binary	69.6%
A racialized person/person of colour	45.6%
LGBTQ	19.0%
A person with disabilities	7.6%
Indigenous/aboriginal	3.8%
None of the above	8.9%

Program Participation

Seventy-eight participants completed the Elements of AI online course. Just under half (46.3%) of survey respondents contributed to Slack discussions during the six weeks allocated to the online course (S1 Q6, n=80).

Fifty participants attended at least four of six of the in-person sessions and 71 participants attended at least one session. An average of 27 participants attended each session in person, and 21 additional people participated by watching a video and contributing to the Slack threads in the month after the live session. The first session had 63 attendees; the final session had 35 attendees.

Forty-eight participants met all graduation requirements. Figure 1.2 shows the number of graduates compared to the original cohort; graduates as a percentage of cohorts were 66.7% (staff), 64.3% (community), 53.3% (faculty), and 38.2% (students).

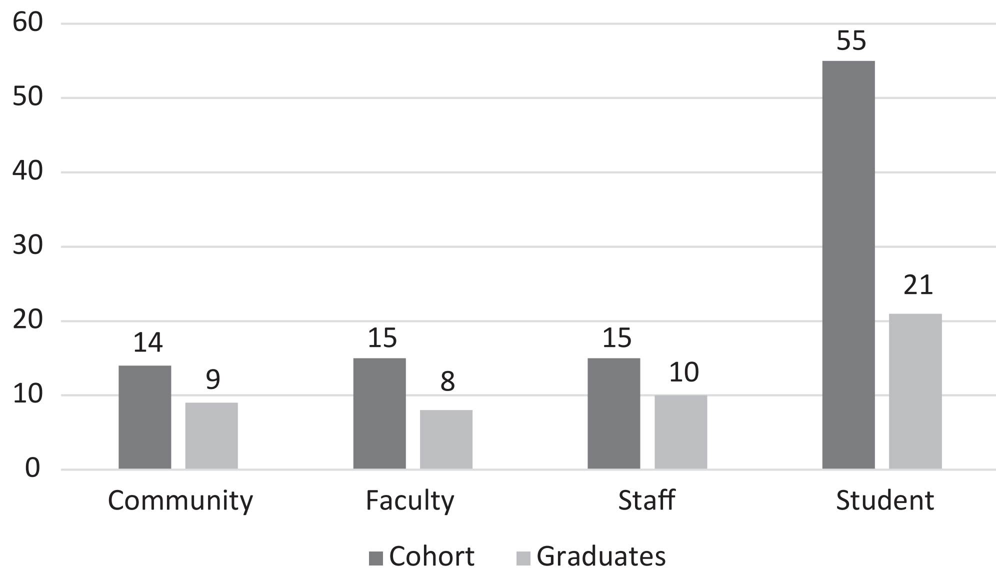


Figure 1.2

Original cohort numbers compared to graduate numbers by status.

Program Evaluation

The feedback from those who did complete the program indicates that it was a success. When asked, "How likely is it that you would recommend the AI Challenge to a friend or colleague?," the average score from respondents was 8.5/10, with almost half of respondents (48.0%) answering with a "10" (S2 Q15, n=50). Four of 51 respondents to Survey 2 said the overall program did not meet their expectations; three of those expected more hands-on practice.

When asked the same question of the Elements of AI OER six months previously, the average score was 8.2/10 (S1 Q12, n=79). Over four-fifths of survey respondents (82.5%) indicated the online course met their expectations (S1 Q3, n=80). Nearly all survey respondents (97.5%) felt their knowledge of AI was stronger or much stronger after the course (S1 Q9, n=80).

While 94.1% of respondents agreed or strongly agreed that the in-person conversations were a worthwhile learning experience (S2 Q3, n=51), some respondents indicated they would like the talks to be longer and for the speakers to consistently stay for the conversation (S2 Q11 & Q24).

Of the 48 respondents who indicated they participated in Slack discussions, 62.5% agreed or strongly agreed that those discussions were a worthwhile learning experience (S2 Q7, n=48), while 22.9% were neutral and 14.6% disagreed or strongly disagreed. One participant noted that "it would be nice to receive feedback on our Slack comments. So much effort went into structuring my thoughts and comments, and I'm not sure they were read by any of the organizers" (S2 Q24).

One way to improve the program could be to further boost engagement between in-person sessions, whether via Slack or other channels; by employing dedicated moderators or by shortening the time between sessions to maintain momentum and minimize attrition, for example, it could run over six months instead of one year. Further development of asynchronous options could also alleviate administrative scheduling and commuting difficulties for participants who primarily attend remotely. More concentrated efforts to engage those who did not complete the program could reveal further insights.

Outcomes

This study sought to identify how The AI Challenge impacted the participants' attitudes and perceptions of AI. In both surveys, participants described how it helped to strengthen their understanding and awareness of AI, to shift their attitudes regarding the technology, to create an interest in learning more about AI along various pathways, and to develop a critical and nuanced attitude toward AI tools and platforms in their personal and professional lives.

Building AI Competency

Most participants did not have a solid understanding of artificial intelligence when they started Phase 1. The first survey sought to capture the impressions of participants after they completed the online course. Over 97% of respondents indicated their knowledge of AI was “stronger” or “much stronger” after the course (S1 Q9, n=80). When asked to “list three key things you have learned in this course” (S1 Q10; n=78), respondents referred to the basic workings of AI (“how AI solve puzzles like GPS routing by decomposing problems in states and transitions”) as well as to applications of the technology (“how AI could be used to create realistic-looking videos and pictures”), implications (“we don’t need to be scared of robots taking over the world”), and limitations of the technology (“while powerful, AI technologies are still quite basic in that they are narrow”). Respondents were also asked to list up to three questions that they still had regarding AI, its applications, or its implications (S1 Q15, n=72). Figure 1.3 shows the contrast between the content of the course and areas where the respondents still had questions; this input informed the development of the conversation series.

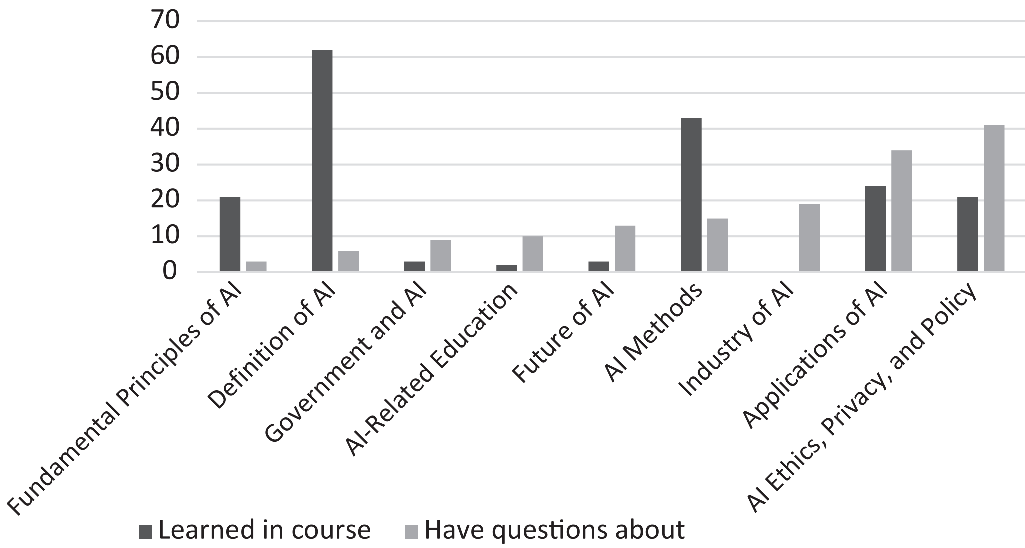


Figure 1.3
Respondents’ areas of learning and areas with questions after Phase 1.

When asked, “How, if at all, have your opinions or assumptions regarding AI changed, based on what you learned in this course?” (S1 Q11), participants felt that they were now able to look at articles about AI more critically and be able to separate the fantasy about AI from its reality, as summarized in this comment: “I thought that AI was beyond comprehensibility in many ways. As a non-technical person, I was intimidated. However, I can now understand its basic tenets, which proves to me that the field, or at least non-technical roles, are accessible to me.”

In the Phase 2 survey, 90% of respondents felt that participating in the challenge increased their knowledge and understanding of AI (Q18, n=49). About three-quarters of participants agreed or strongly agreed that they felt confident about understanding, explaining, and examining aspects of AI (S2 Q19). When asked how their perspective on

AI specifically changed since joining the AI Challenge (S2 Q20), coding revealed a general theme toward broadened, deepened understanding. In the words of one respondent, "AI overall has become less of a 'black box,'" while another said they have "come to… appreciate its benefits and biases." Several explained how the course brought AI out of science fiction into the present reality: "From Google Maps to Uber… we can't be worried about AI coming because it's already here in so many ways." Others reflected on the impact of media on their perceptions coming into the course: "*Terminator* is not a good starting point for AI." For some, AI had become more visible: "I had not appreciated the pervasiveness of the existence of AI nor appreciated how hidden much of it is." For others, the limitations were now in clearer focus: "Now I understand AI mainly as a new form of tool—potentially transformative or even revolutionary, but, like all tools, only as good as the people using it."

Perspective Shift

The change in participants' attitudes regarding AI can be characterized as a shift toward ambivalence, defined here as feeling both positively and negatively about AI for different reasons. On one hand, demystification of the technology often reduced anxiety: "I've been given a more complex way of viewing AI through the discussions in this challenge that has largely removed my fear of it" (S2 Q20). On the other hand, a greater understanding of the impacts and implications of technology on vulnerable populations was a cause for concern for some: "I was unaware how terrible systems that pre-exist are exacerbated or made worse by AI," "knowing how AI can be misused has me feeling a bit uneasy." For another participant, the live sessions gave words to what was before more intuitive or abstract: "I still have concerns about AI, but my formerly vague fears have become more specific and articulate, and I now feel that I can discuss them intelligently and move a discussion forward instead of just reacting."

When asked to gauge their "primary attitude or position on AI" coming into Phase 1 of the program in summer 2019 (S2 Q16) compared to at the completion of Phase 2, accounting for what they learned over the course of the challenge (S2 Q17), participants did not express a uniform shift. Answer choices ranged from "largely positive" to "largely negative," with options for "unsure" ("I don't understand AI enough to make a decision"), "ambivalent" ("I feel both positively and negatively about AI for different reasons"), and "neutral" ("I feel neither positively nor negatively about AI"). Nine of 10 people who felt largely negative before now felt ambivalent (6) or largely positive (3), with one person remaining in the "largely negative" group. That no one felt "unsure" afterward, versus 8 of 50 participants before, suggests that everyone had learned enough to form an opinion. Thirty-one people felt "ambivalent" afterward, showing critical thinking learned through the program (compared to 13 people previously). As one person explained (S2 Q22), "although my answer is… 'Ambivalent' in both cases—that doesn't mean my perspective on AI hasn't changed.… I learned so much about what questions to ask when evaluating the use of AI." This ambiguity can also be seen in participants' responses when asked to rate their confidence in using AI technologies in their personal and professional lives (S2 Q19). Respondents indicated they were less confident overall in using AI than they were in understanding, explaining, and examining AI. "Overall, I would say I have gone from being concerned about AI technologies to being excited yet wary."

AI in Application

The cautious attitude that many participants describe translated into recommendations for regulation, oversight, and education when implementing AI technologies (S2 Q22). While participants saw benefits of AI applications such as processing huge volumes of data, they rejected its use in instances "wherever empathy and emotion are necessary ingredients (immigration, medicine)." Another recurring theme was the "glaringly, repeatedly obvious" need for diversity in both data and among those who analyze it. Not surprisingly, the call for more education was also a theme: "Society must understand how these tools are being used so they can make informed decisions and be critical of information presented to them."

Participants were enthusiastic about continuing to learn about AI. When asked how they have or will apply the knowledge/skills acquired, responses ranged from teaching, classroom learning, and research projects to less formal learning opportunities or even switching jobs or fields. One faculty member said, "I have already incorporated some of the concepts into the curriculum of one of my courses and am using the insights to help re-imagine the competencies that professional [redacted] should have" (S2 Q14). These areas are of particular interest to librarians supporting faculty and students in academic institutions.

Conclusion

The 99 AI Challenge gave a cross-section of the U of T community the opportunity to engage with resources and support that they might not otherwise have been able to access. The historical lack of diversity in technical fields such as computer science was identified by the organizers at the outset of this project as an opportunity to enhance the collective knowledge at U of T in terms of AI and related issues. The participants also highlighted this need: one of the program's greatest strengths was that it brought together people of diverse backgrounds to discuss issues relevant to all, and having these voices is important for the responsible development and use of AI applications. The search for knowledge and understanding of AI and its applications, benefits, ethical dilemmas, and drawbacks is one that shouldn't be limited to technical experts or STEM disciplines.

Libraries, as trusted providers and partners in open community-based learning, are well-positioned to lead this exploration. As faculty and researchers increasingly employ AI in teaching and research contexts, and as industry and government increasingly explore and use AI, librarians stand to benefit from increased capacity and understanding as educators and as users of these technologies in library systems.

Endnotes

1. Chris Sorenson, "How U of T's 'Godfather' of Deep Learning Is Reimagining AI," U of T News, November 2, 2017, https://www.utoronto.ca/news/how-u-t-s-godfather-deep-learning-reimagining-ai.

2. Janosch Delcker, "Finland's Grand AI Experiment," Politico, January 2, 2019, https://www.politico.eu/article/finland-one-percent-ai-artificial-intelligence-courses-learning-training/.
3. Ashok K. Goel, "AI Education for the World," *AI Magazine* 38, no. 2 (Summer 2017): 3, https://doi.org/10.1609/aimag.v38i2.2740.
4. "Table 322.40. Bachelor's Degrees Conferred to Males by Postsecondary Institutions, by Race/Ethnicity and Field of Study: 2016–17 and 2017–18," National Centre for Education Statistics, *Digest of Education Statistics* (October 2019), https://nces.ed.gov/programs/digest/d19/tables/dt19_322.40.asp.
5. "Table: 37-10-0011-01: Postsecondary Enrolments, by Field of Study, Registration Status, Program Type, Credential Type and Gender," Statistics Canada, 2020, https://doi.org/10.25318/3710001101-eng.
6. Natasha Singer, "The Hardest Part of Computer Science? Getting into Class," *New York Times* (January 24, 2019), https://www.nytimes.com/2019/01/24/technology/computer-science-courses-college.html.
7. "Diversity and Inclusion Report 2019," Microsoft, 2019, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4aqv1.
8. "Tech-Founder Education: Backgrounds of 585 Founders from 335 Canadian Tech Companies," Impact Centre, University of Toronto, 2019, https://www.impactcentre.ca/wp-content/uploads/2019/02/Tech-Founder-Education.pdf.
9. Michael Wollowski et al., "A Survey of Current Practice and Teaching of AI," in *Thirtieth AAAI Conference on Artificial Intelligence* (Phoenix: AAAI Publications, 2016), 4121–22, https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12444.
10. Judy Goldsmith and Emanuelle Burton, "Why Teaching Ethics to AI Practitioners Is Important," in *Workshops of the Thirty-First AAAI Conference on Artificial Intelligence* (San Francisco: AAAI Publications, 2017): 4538, https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/download/14271/13992&sa=U&ved=2ahUKEwiVorb3paTzAhV9MlkFHfBSCjAQFnoECAUQAQ&usg=AOvVaw2_jN7K56cXho9PCyYzPeKZ.
11. Goel, "AI Education," 3.
12. Ibid.
13. Amanda Wheatley and Sandy Hervieux, "Artificial Intelligence in Academic Libraries: An Environmental Scan," *Information Services & Use* 39, no. 4 (February 6, 2020): 353, https://doi.org/10.3233/ISU-190065.
14. Wheatley and Hervieux, "Artificial Intelligence in Academic Libraries," 354.
15. Kenning Arlitsch and Bruce Newell, "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries," *Journal of Library Administration* 57, no. 7 (October 3, 2017): 795, https://doi.org/10.1080/01930826.2017.1362912.
16. "Artificial Intelligence and Intellectual Freedom, Key Policy Concerns for Canadian Libraries," Canadian Federation of Library Associations, July 2018: 5–6, http://cfla-fcab.ca/wp-content/uploads/2018/07/CFLA-FCAB-2018-National-Forum-Paper-final.pdf.
17. "Artificial Intelligence," American Library Association, *Centre for the Future of Libraries—Trends*, February 4, 2019, http://www.ala.org/tools/future/trends/artificialintelligence/.
18. "Urban Libraries Take a Leading Stance to Advance Equitable Data Governance and Literacy in the Age of Artificial Intelligence," Urban Libraries Council, January 16, 2019, https://www.urbanlibraries.org/files/AI-Working-Group-Press-Release.pdf.
19. Larry Alford, E-Mail to the Authors, February 3, 2019.
20. "Careers at U of T," Human Resources & Equity, n.d., https://careers.hrandequity.utoronto.ca/.
21. "Elements of AI," Reaktor, 2020, https://www.reaktor.com/elements-of-ai/.

Bibliography

American Library Association. "Artificial Intelligence." Centre for the Future of Libraries—Trends. February 4, 2019. http://www.ala.org/tools/future/trends/artificialintelligence.

Arlitsch, Kenning, and Bruce Newell. "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries." *Journal of Library Administration* 57, no. 7 (October 3, 2017): 789–98. https://doi.org/10.1080/01930826.2017.1362912.

Canadian Federation of Library Associations. “Artificial Intelligence and Intellectual Freedom, Key Policy Concerns for Canadian Libraries.” July 2018. http://cfla-fcab.ca/wp-content/uploads/2018/07/CFLA-FCAB-2018-National-Forum-Paper-final.pdf.

Delcker, Janosch. “Finland’s Grand AI Experiment.” Politico. January 2, 2019. https://www.politico.eu/article/finland-one-percent-ai-artificial-intelligence-courses-learning-training/.

Goel, Ashok K. “Editorial: AI Education for the World.” *AI Magazine* 38, no. 2 (Summer 2017): 3–4. https://doi.org/10.1609/aimag.v38i2.2740.

Goldsmith, Judy, and Emanuelle Burton. “Why Teaching Ethics to AI Practitioners Is Important.” In *Workshops of the Thirty-First AAAI Conference on Artificial Intelligence*. San Francisco: AAAI Publications, 2017. https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/download/14271/13992&sa=U&ved=2ahUKEwiVorb3paTzAhV9MlkFHfBSCjAQFnoECAUQAQ&usg=AOvVaw2_jN7K56cXho9PCyYzPeKZ.

Human Resources & Equity. “Careers at U of T.” N.d. https://careers.hrandequity.utoronto.ca/.

Impact Centre. “Tech-Founder Education: Backgrounds of 585 Founders from 335 Canadian Tech Companies.” University of Toronto. 2019. https://www.impactcentre.ca/wp-content/uploads/2019/02/Tech-Founder-Education.pdf.

Microsoft. “Diversity and Inclusion Report 2019.” 2019. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4aqv1.

National Centre for Education Statistics. “Table 322.40. Bachelor’s Degrees Conferred to Males by Postsecondary Institutions, by Race/Ethnicity and Field of Study: 2016–17 and 2017–18.” *Digest of Education Statistics*. National Center for Education Statistics. October 2019. https://nces.ed.gov/programs/digest/d16/tables/dt16_322.50.asp?current=yes.

Reaktor. “Elements of AI.” 2020. https://www.reaktor.com/elements-of-ai/.

Singer, Natasha. “The Hardest Part of Computer Science? Getting into Class.” *New York Times* (January 24, 2019). https://www.nytimes.com/2019/01/24/technology/computer-science-courses-college.html.

Sorenson, Chris. “How U of T’s ‘Godfather’ of Deep Learning Is Reimagining AI.” U of T News. November 2, 2017. https://www.utoronto.ca/news/how-u-t-s-godfather-deep-learning-reimagining-ai.

Statistics Canada. “Table: 37-10-0011-01: Postsecondary Enrolments, by Field of Study, Registration Status, Program Type, Credential Type and Gender.” 2020. https://doi.org/10.25318/3710001101-eng.

Urban Libraries Council. “Urban Libraries Take a Leading Stance to Advance Equitable Data Governance and Literacy in the Age of Artificial Intelligence.” January 16, 2019. https://www.urbanlibraries.org/files/AI-Working-Group-Press-Release.pdf.

Wheatley, Amanda, and Sandy Hervieux. “Artificial Intelligence in Academic Libraries: An Environmental Scan.” *Information Services & Use* 39, no. 4 (February 6, 2020): 347–56. https://doi.org/10.3233/ISU-190065.

Wollowski, Michael, Robert Selkowitz, Laura E. Brown, Ashok Goel, George Luger, Jim Marshall, Andrew Neel, Todd Neller, and Peter Norvig. “A Survey of Current Practice and Teaching of AI.” In *Thirtieth AAAI Conference on Artificial Intelligence*. Phoenix: AAAI Publications, 2016. https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12444.

Chapter 2

URI Libraries' AI Lab—

Evolving to Meet the Needs of Students and Research Communities

Harrison Dekker, Angelica G. Ferria, and Indrani Mandal

Introduction

In 2018, the University of Rhode Island launched what was most likely the world's first library-based artificial intelligence (AI) lab. While an AI Lab may seem incongruous with traditional perceptions of the library, it is actually consistent with a growing trend of libraries providing space and support for students, faculty, staff, and communities engaged in computational research. Across the literature, new job titles such as data librarian, emerging technologies librarian, and reproducibility librarian are emerging to provide support for data-intensive research.[1] Depending on the desired level of service a library intends to provide, these positions may demand domain expertise and technical skills in programming and statistical analysis that are not traditionally associated with librarianship.[2]

Behind this trend are the same technological factors that are driving rapid change across society—namely, faster computers and networks, open software, and research practices that utilize ever-expanding amounts of data. Just as libraries have traditionally provided support to researchers requiring guidance in the often complex scholarly publishing landscape, modern library professionals are guiding researchers through new domains associated with data-centric research.[3] Given the rapid pace of change, and the inability of the academic curriculum to keep pace, librarians are finding new roles as trainers and advisors to researchers who have high levels of domain knowledge but less experience with programming platforms.

But how farfetched is it for librarians to be taking on these new roles? Considering the interrelationship between scientific research and the open source software movement, clues can be found for why librarians are finding an affinity for these roles. For instance, building, maintaining, and navigating the evolving ecosystem of open source software tools parallels the traditional librarian ability to select, manage, and search databases and collections. In the past, when a researcher wanted to utilize an analytical method not provided by any commercially available software, it was usually necessary to build a new software tool from scratch, which, depending on the complexity of the task, might require a significant investment in programming expertise. Now, free and open source programming platforms allow new functionalities to be added to existing software packages with significantly less investment. In conjunction with this trend has been the evolution of trusted distribution networks such as CPAN,[*] CRAN,[†] and PyPI,[‡] which facilitate sharing, discovery, and reuse of community-developed software tools. As a result, assisting researchers in finding and selecting the right tool for the job in a constantly shifting landscape is not unlike the service librarians provide in helping users navigate more conventional repositories of information.[4] Similarly, traditional librarian knowledge of search, information organization, and provenance all come into play.[5] Moreover, the traits expressed in the open source software environment—free or nominal cost, easily and reliably shared, and continuous development with version control—would thrill libraries if adopted by publishers.

Accompanying this evolution of library services has been a parallel trend of libraries offering dedicated *maker* and *data lab* spaces to facilitate access to software, equipment, and expertise that fall outside their traditional purview.[6] Justifications for this phenomenon have been explored in the literature. John Burke makes the case that "if an academic library can commit time, space, and a little money, and serves a campus community that is interested in exploring experiential learning, a makerspace and making programs can be built and can thrive." His argument is based on two distinct justifications, the first having to do with the demand for and efficacy of experiential learning and the second revolving around the inherent suitability of the library environment for delivering these opportunities.[7] Building upon these ideas, "A Studio Model for Academic Data Services" uses the metaphor of an academic art or design studio as a framework to explore the characteristics of successful library data service spaces.[8] The key ingredients include staff with specialized expertise, resources, space, and a learning community engaged in creative, iterative, and self-directed work.[9]

The URI Library AI Lab extends this data studio model by focusing not just on data and code but also on AI technology, like programmable robots, high-end laptops (TensorBook[§]), and access to high performance computing (HPC) resources. This equipment is

* The Comprehensive Perl Archive Network(CPAN), see: www.cpan.org.

† The Comprehensive R Archive Network (CRAN), see: https://cran.r-project.org/.

‡ The Python Package Index (PyPI), see: https://pypi.org.

§ TensorBook, see: https://lambdalabs.com/deep-learning/laptops/tensorbook.

available to any student who chooses to use the lab, whether for a course assignment or to simply explore an interest outside their major field. Positioned in a prominent location in the library lobby and promoted as a place where all are welcome, the AI Lab is intended to create a space where students feel free to collaborate regardless of background to address the challenges of AI, which are, after all, fundamentally interdisciplinary.

History and Funding

The URI University Libraries' Artificial Intelligence Lab (AI Lab) is a product of a collaboration between multiple URI colleges, academic groups, and university technology services. Working with partners, URI Libraries submitted a proposal to the Champlin Foundation, a local organization with a long-standing history of supporting the University of Rhode Island in technology projects that deliver a positive impact to Rhode Island residents.

A group consisting of professors from the College of Engineering (biomedical) and the College of Arts and Sciences (philosophy), the head of the university's Big Data Initiative, and the library's data librarian formed the core AI Lab team. Envisioning "an easy-to-access facility at a centralized location," the AI Lab team described "an information rich source for those wishing to learn about artificial intelligence both theoretically and practically"—the first library-based lab of its kind in the world.

Invited to present the project to the Champlin Foundation Committee, the team drew attention to the joint nature of the endeavor, citing the $60,000 in contributions from the College of Engineering, the Provost's Office, and the library had already set aside to cover the cost of updating wiring, carpet, paint, and furnishing in the Lab's intended location. Supported by the in-kind contributions, the AI Lab team could assure the Foundation Committee that the amount requested in the proposal, an additional $180,000, would go directly to the purchase of equipment.

First Steps

URI's University Libraries AI Lab celebrated its grand opening in September 2018. Behind the scenes, much of the Lab's equipment sat unassembled. In addition, the Lab's most important piece of equipment, the NVIDIA DGX-1 server, had yet to be installed; it sat in the University's Advanced Computing Cluster (a group of servers dedicated to research computing) awaiting a cable that had not been ordered.

Fortunately, enthusiasm for the project, evident as students, faculty, and university administrators detailed the value of the resource, exceeded any setbacks. The AI team shared the Lab's vision with an audience aware of the academic potential. Drawing attention to the location within the library, the team highlighted the Lab's purpose, encouraging cross-discipline collaboration and fostering understanding to ensure responsible development and adoption of future technologies.

The library hired a part-time lab manager, a machine learning (ML) expert also employed at URI as a lecturer in the Computer Science and Statistics Department, as well as several student workers. The lab manager assigned student workers projects intended to complement their interests. The first tasks completed involved robot assembly from simple "toys" to more advanced machines requiring an understanding of robotics and coding. While the robots attracted the attention of the curious, the first actual users were students seeking advanced computing resources that, without the library's AI Lab, would not have been available. Given the steep learning curve inherent in utilizing these resources, the development of tutorials and learning modules quickly became a priority.

Purchasing

Due to restrictions from both the Champlin Foundation and the university, the AI Lab was planned and proposed with no financial requests for staffing. The amount from Champlin was to go to technology, while the in-kind contributions from the university would provide the minimal wiring, some new carpet and paint, and the furnishings required to transform an outdated government publications office into a small yet engaging AI Lab.

It is a requirement of the Champlin Foundation that funds supporting technology projects be entirely spent within the awarded calendar year. Informed of success in October of 2017, with monies to be dispersed the upcoming January, the library immediately began spending the in-kind contributions, placing necessary work orders and contacting the university's primary furniture vendor.

In recognition of the time requirements of university and State of Rhode Island procurement processes, contact was made with potential vendors, requesting quotes and encouraging those expressing interest to register as both State of Rhode Island and (as a separate process that never failed to cause confusion) university vendors. While the authors' personal experiences regarding the intricacies of procurement processes could easily hijack the rest of the chapter, for the sake of mutual sanity, it is sufficient to say the technology purchases were completed by the December deadline.

Equipment

At the time of writing, the technology purchased with the Champlin Foundation funds in 2018 has met the AI Lab's needs with few additions. Figure 2.1 provides a basic idea of the AI Lab's original equipment.

Engagement

In tandem with establishing a lab space for student skill development, the AI Lab is also committed to creating an intellectual space for the exploration of conceptual and cultural issues related to the development of artificial intelligence today. Within this "space," the library convenes conversations at various levels within and beyond the URI community, first by holding events to identify and bring together faculty and staff with an active interest in AI from diverse vantage points, and second by generating curricular products and public programming that grow out of identified interests. After all, AI is a very timely topic in which global entrepreneurs, scientists, and leaders weigh in on its potential benefits and harms.

Computing	Collaboration
Nvidia DGNvidia DGX-1 (GPU Server) Lambda Tensorbooks (Lab Workstations)	Samsung Flip Digital Display NAS Hard Drives
Robots	**Internet of Things (IoT)**
Robotis Turtlebot 3 Burger and Waffle EZ-Robot JD Humanoid Robot EZ-Robot Revolution Six WiFi Hexapod Vex IQ Super Kit Vex IQ Foundation Motion Add-On Kit	Smartwatch Amazon Echo Apple Home Xbox controller Nighthawk XR (router) Jetson RX2 Development Kit
List does not include peripherals (keyboards, monitors) and consumables (batteries, IoT kits, sensors, etc.).	

Figure 2.1

AI Lab Equipment Table

Discussion of AI, no matter the level of technological complexity, includes exploration of potential impacts, be they ethical, societal, philosophical, or economic. Leveraging the library as a crossroads of knowledge, the flexibility of AI as a topic for themed programming, special events, and community outreach is limitless. The URI AI Lab has utilized AI's novice-to-expert span of appeal to develop, deliver, or host everything from AI Summer Camps for students from disadvantaged secondary schools to partnering with NVIDIA to deliver a Fundamentals of Deep Learning in Computer Vision workshop to graduate students.[10]

While a more comprehensive list of AI Lab activities may be found in appendix 2A. A few are highlighted below.

University Libraries' Artificial Intelligence Summer Camps

Beginning in summer 2019 (at the time of writing, the AI Lab had finished delivering the 2020 camps), the AI Lab has run a series of week-long, full-day summer camps for Rhode Island secondary school students. While it is important to note the AI Lab actively pursues, encourages, and supports diversity, inclusion, and equity at every opportunity, the manner in which the AI Summer Camps in particular are offered provides insight into the Lab's commitment to inclusivity.

The original AI Summer Camp offered in 2019 charged campers a rate similar to URI's other Summer Camp offerings. Limited to fifteen students, these face-to-face camps delivered AI Lab-designed content and curriculum taught by URI graduate students to Rhode Island students in grades three to twelve. Finishing the five-week series in July, the AI Lab partnered with the College of Engineering's Diversity Office to run an additional camp for students from underprivileged Rhode Island schools. Delivering this free camp allowed the AI Lab to engage with local schools while exploring the feasibility of offering camps free or at a reduced cost.

In February 2020, the AI Lab began planning for summer. Knowing the operational costs and ability to re-use many of the robots, kits, and supplies of the previous year, it was decided the cost of the camps would be reduced, an early-registration discount would be offered, and a few seats in each camp would be reserved for scholarship campers (application process not entirely determined). That plan, rough sketch that it was, was dropped as the COVID-19 pandemic took hold. By the beginning of May, it was apparent that the suspension of face-to-face and group activities would continue into the summer.

The AI Lab decided to embrace society's mass online migration and run virtual Summer Camps. Redesigning the curriculum of both previous and upcoming camps (those that could be adapted) to an online format, a schedule of full-day virtual camps was created. Recognizing that the new camps could not offer the day-camp experience and yet were much more accessible, it was decided the camps would be free (the fifteen-camper limit was maintained to ensure the quality of instruction would remain). Preference was given to applicants who self-identified as students of under-privileged Rhode Island schools.

Exploring Evidence: Ethics and AI Workshop Series

Scheduled for spring 2020, the AI Lab partnered with a lecturer in the Philosophy Department to offer three interactive workshops exploring the intersection of artificial intelligence, society, and ethics. Each hour-and-a-half in-person workshop would introduce attendees (limited to twenty individuals) to the increased potential for human rights violations and discrimination engendered by the unethical application of AI technology. In addition, these workshops were designed in a manner that would allow the AI Lab to offer one, or all, again at a future date with an AI Lab graduate student or librarian as instructor.

OLLI Artificial Intelligence: The Game Changer

One of the AI Lab's more surprising audiences, the Osher Lifelong Learning Institute (OLLI) at URI, has expressed continued interest in AI Lab operations and events. Frequent attendees of the AI Lab and AI-RI Meetups, the OLLI members requested that the AI Lab present to their group in fall 2019. Accompanied by two student workers, members of the AI Lab team shared the Lab's mission, goals, and programs supplemented with slides detailing Lab equipment and technical specifications. After providing a basic introduction to machine learning and artificial intelligence, AI Lab student workers stepped in, presenting their projects and explaining how the work related to their classes and future career prospects.

In the resulting thank you, the organizer expressed the group's appreciation for the AI Lab team's willingness to stay beyond the scheduled presentation time and for defining AI-associated acronyms, mentioning in particular the benefit of having NLP (natural language processing) explained, as it had been a reoccurring topic in the news at the time. After commenting on the growing ubiquity of AI, the note emphasized the value of the students' presentations "…as they are soon to be in the AI workforce and contributing to the well-being of those who did not grow up digital natives."

The Vision Meets Reality

The necessity of operating within constraints has shaped an AI Lab that embodies the vision if not the structure presented in the proposal. The original proposal described the AI Lab as two very separate entities: a physical space and an incorporated space where "creative ideation around Artificial Intelligence" could occur. In practice, the AI Lab is much more of an advanced technology access point, guided by the theory that to advance understanding of AI in all forms, the learners should be met where they are.

Two critical factors impacted how the project evolved. First, in contrast to the collaborative nature of the original proposal, deployment was delegated entirely to the library. This meant that library staff handled the selection and procurement of actual equipment and that ultimate staffing decisions were directed by the library. Given the low level of domain expertise within the library, it was necessary to work with partners to recruit hourly staff with sufficient expertise to fill the gaps. Another decision made on the fly was to realign the position of the recently hired data librarian, who had relevant programming expertise as well as experience with establishing a library-based lab for data analysis.

The second factor critical to understanding the priorities that emerged in the rollout of the AI Lab was that the librarians involved were actual AI novices. As exciting as it would have been to have the lab full of technology experts building autonomous robots and the like, the reality was that most users needed to focus on the basics, like learning to program, which was something the Lab was prepared to teach. In this regard, the staff's status as AI novices actually worked to their advantage in planning and implementing the new service.

It would have been easy to pigeonhole the Lab, focusing primarily on the high-tech equipment comprehensible to only a small subset of the community and foregoing the problems involved in making the Lab accessible to any user regardless of prior experience with AI technology. However, such a decision would be in conflict with the library's commitment to accessibility, equity, and inclusion, values that had been made clear to stakeholders throughout the project proposal. As learners themselves, the small team behind the AI Lab could explore topics, projects, partnerships, technologies, and programs through a broad lens, ensuring the Lab could adapt to the interests and needs of multiple groups.

The Lab was simultaneously trying to engage a community ready and able to tackle topics such as "What is human?" and develop partnerships to build a framework of workshops and resources to introduce beginners to complicated technology equipment and methods. Luckily, libraries and librarians are adept at supporting users in a way that is consistent with their individual level of expertise.

Lessoned Learned

Staffing

A fundamental challenge in establishing the AI Lab was how to address the imbalance between the rich collection of AI technology afforded through the initial grant with the library's comparatively limited staffing options. This is, of course, a familiar library scenario.

The library was fortunate to have a librarian on staff who, while only able to offer an interim commitment to working with the Lab, had the expertise to teach programming workshops and consult on high performance computing needs. The Lab was even more fortunate to find a part-time lecturer from the Computer Science and Statistics program with expertise in machine learning and a strong commitment to undergraduate as well as K-12-level instruction. A key factor in keeping the individuals in these positions engaged was to give them nearly complete autonomy in how they chose to invest their limited time, accepting that many decisions would be made on the fly and by not setting arbitrary assessment metrics.

This culture of self-direction was encouraged among the student workers who were given opportunities to play with the technology as they assembled new equipment or assisted visitors. Given the amount of enthusiasm the first batch of employees showed for the Lab, this approach seemed to work well during the initial start-up phase as evidenced by their willingness to invite fellow students to the Lab, take part in volunteer activities, or participate in off-campus events.

One unanticipated phenomenon was the number of students interested in doing volunteer work. In hindsight, presented with the ubiquitous predictions surrounding AI's impact on the future job market, this should not have come as a surprise, since it's natural for students to seek applicable experience. For the most part, the Lab avoided creating formal volunteer positions, since students can freely access the Lab and its equipment, but this "non-policy" may be revisited in the future.

Technological Demands

The AI Lab received its startup funding in the form of a grant for the purchase of equipment. Much of this equipment came unassembled, and some items turned out to be highly susceptible to damage or, in a few cases, defective on arrival. The most cost-intensive purchases, namely the DGX-1 server and TensorBooks, required significant setup as well as ongoing system administration and software and hardware maintenance. Student workers assemble, maintain, and troubleshoot most of the non-computer equipment, and the Lab has been fortunate to find and hire students with the requisite skills.

Setup and administration of the TensorBooks were also done in-house, with the exception of a collaboration with URI's IT Research Computing Service in which a portal for remote access was created. The Lab has also relied on the campus IT department to host, maintain, and troubleshoot the DGX-1 server and provide occasional technical support to users, although the Lab is capable of providing training and support to onboard new users. It must be pointed out, however, that many of the users have never worked in a high performance computing environment and, therefore, in almost all cases, require assistance. The biggest barrier for users thus far is the requirement that their AI code is run within a container environment, which in this case is an open source tool called Singularity.* Containers are a critical tool to solve the system administration challenge of creating a unique computational environment that meets the needs of individual users without impacting the environments of other users or the base system.[11] In scientific computing, this is particularly important, since there is rarely a one-size-fits-all solution, due to interdependencies between the various programming components.

* Singularity, see: https://sylabs.io/singularity/.

Community Building

By providing access to equipment previously unavailable to students and faculty, the AI Lab has laid a foundation for fundamentally changing how teaching and learning about AI technology and issues is carried out at URI. One critical impact that has remained elusive, however, is the integration of the Lab's resources into the curriculum. While some students do utilize the Lab on their own initiative, the reality is that for most students, particularly in STEM disciplines, the curriculum does not afford them much spare time. Faculty routinely express interest in modifying their syllabi to introduce an assignment that utilizes the Lab, but it is a lot of work for them to take this on, particularly if they are unsure of the outcome. One success on this front was the inclusion of an assignment in an undergraduate class that required the students to make use of the high-end laptops. Ideally, more faculty will make the leap and adapt their courses to take advantage of lab resources in the future. However, given the steep learning curve of our AI resources and the amount of hands-on assistance required, it may be the case that one-off student projects, such as capstone projects and independent study, will have the most impact in this regard. That said, if collaboration with external departments remains elusive, the library will develop and teach its own for-credit courses.

An important step taken by the library toward community-building has been the creation of a faculty position shared between the library and the Department of Computer Science and Statistics. A key responsibility of this position in addition to teaching will be to pursue funding opportunities to help support AI Lab activities.

Conclusion

The growth in demand for AI resources and training that motivated the creation of the AI Lab is part of a broader technological trend that allows for the creation, transmission, and processing of ever-expanding amounts of data. Given the popularity and demand for the AI Lab's programming workshops and ongoing discussions about a campus need for a centralized interdisciplinary data consulting and training service, it seems likely that there will be some sort of convergence. If that does occur and the AI Lab is able to offload some of its current offerings, the future is likely to more prominently feature some combination of the following:

- continuation of the speaker series, with a focus on emerging social, ethical, and professional topics
- focus on AI-specific equipment in support of STEM education, both for URI students and the K-12 programs, and a continued push for new or modified courses to ensure continuous and predictable demand for services (This trajectory would be enhanced by the creation of a position for someone with the requisite STEM background to mentor students and collaborate with faculty.)
- development of a research agenda around the application of AI technologies, such as natural language processing or certain forms of image processing that are aligned with specific library needs (If successful, this could also help position the library as a campus center for expertise in these types of AI applications and strengthen its role as a research partner across disciplines.)

Appendix 2A
AI Lab Workshops

Machine Learning
Fundamentals of Deep Learning for Computer Vision—NVIDIA Deep Learning Institute.

Presented in partnership with the Department of Computer Science and Statistics and NVIDIA, this workshop teaches deep learning techniques for a range of computer vision tasks through a series of hands-on exercises.

Natural Language Processing (NLP)
Learn two methods of NLP—text mining and topic modeling—through hands-on R programming practice.

Machine Learning Boot Camp
Advanced Machine Learning for research and coursework.

Intro to Machine Learning
Build machine learning models with us through WebEx sessions. Integrate machine learning into your research.

Data Science

Introduction to Python
Python for Data science focuses on those researchers who have been using R or SAS for their data science research. Explore Python with us and integrate it in your research.

- Data cleaning—Pandas
- Plotting—Matplot, NumPy
- Database programming in Python

Python
Four online workshops on Python, Python for Data Science and Machine Learning. We will work on some code together and learn how to integrate Python in your research.
Introduction to Python

- Lists, loops, and functions
- Classes and objects
- Inheritance/Dictionary
- Introduction to ROS

R

- Introduction to R
- R for the Social Sciences
- Data Carpentry Bootcamps

GitHub

- Building a Professional Portfolio with GitHub
- Use the popular code-sharing platform GitHub to showcase samples of your academic or professional work.

Appendix 2B
Meetup Events

Inaugural Rhode Island AI Meet-up!

February 23, 2018	Inaugural Rhode Island AI Meet-up!	Karim Boughida, the dean of the URI Libraries will give a short introduction about the new AI Lab, which is slated to open in the fall of 2018. Check out the news article about the new AI Lab here: https://www.insidehighered.com/news/2018/01/17/rhode-island-hopes-putting-artificial-intelligence-lab-library-will-expand-ais-reach. Free discussion to follow afterward. This will be our first meeting! We will welcome everyone, share our interests in AI, and talk about what topics we want to discuss at future meet-ups.
March 30, 2018	Rhode Island AI Meet-up: building AI lab programs	This will be our 2nd meeting! We will welcome everyone to share ideas on how to build a community-based AI lab at URI.
April 4, 2018	#POCAI18 (People of Color-AI) Rhode Island AI Meet-up	This is our 3rd meetup. Join us to discuss ethical considerations underlying AI and machine learning projects and how diversity and inclusion are handled in the field of AI. We welcome all of the URI Community and beyond; this is an interdisciplinary event. A light lunch will be served: Program: 12:30 pm–3:00 pm: Introduction: Donald Dehayes, Provost and Vice President for Academic Affairs, URI Welcoming remarks: Naomi Thompson, Associate Vice President and Chief Diversity Officer, URI Moderator: Karim Boughida, Dean of University Libraries, URI Presenters: Dr. Timnit Gebru: researcher at Microsoft Research, New York City in the FATE (Fairness Transparency Accountability and Ethics in AI) group. She was a PhD student in the Stanford Artificial Intelligence Laboratory and Co-founder of the group Black in AI. Dr. Ahmed Bouzid: Co-founder and CEO at Witlingo. Previously, Head of Product with Amazon.com's Alexa/Echo group, and earlier VP of Strategy & Innovation at Genesys. This event is sponsored by the URI Office of Community, Equity and Diversity and the Multicultural Student Services Center and a partnership of URI Libraries Big Data Collaborative and Diversity Initiatives.

December 4, 2018	Ethics and Artificial Intelligence	Join us in the URI Library for an update on the AI Lab's first semester, followed by an open conversation on Ethics and Artificial Intelligence with Doug Friedman, Data Science Manager (Johnson & Johnson), Associate Prof. Harrison Dekker (URI Libraries), and Prof. Cheryl Foster (URI Department of Philosophy. Moderated by Prof. Joan Peckham and convened by AI Lab Instructor, Indrani Mandal. Light fare provided.
March 4, 2019	Robots with Legs	About Yeuhi Abe: Yeuhi Abe is a Senior Roboticist at Boston Dynamics. He received his PhD on physics-based methods for animating humanoid characters at MIT. About the talk: As our society becomes accustomed to automation in the realm of digital information, many predict an extension to the physical world. As robots will become a common tool, they will wander our streets, enter our homes, and help maintain our wild lands. What physical characteristics will these robots take and how will they move? Legs are a nature-inspired solution, fit for go-anywhere mobility. This talk will discuss the advantages and challenges of legged robots, with a focus on Boston Dynamics' history of legged robot development.
March 22, 2019	Bias, Jobs and the Future of AI and Robotics / Peter Haas	Should we be scared of robots and AI? Peter Haas Associate Director of the Humanity Centered Robotics Initiative (Brown University) touches on the social impacts AI and Robotics will have in the next ten years. He explores topics of algorithmic bias and automation-driven job displacement, painting a future picture that is simultaneously grim and hopeful.
April 11, 2019	The Future of Work / Darrell West: Author Lecture and Book Signing	From automation and digital economies to health care and life-long learning, "The Future of Work" explores possible solutions to the social, economic, and political challenges facing society as the concept of "work" is redefined. Join the RI-AI Meetup and Darrell West, founding director of the Center for Technology Innovation at Brookings and Editor-in-Chief of TechTank, as we discuss "The Future of Work." *The Future of Work: Robots, AI and Automation* can be found at the URI Campus Bookstore and available during the event. Hardcover: 9780815732938; Ebook: 9780815732945.

April 30, 2019	Artificial Intelligence: American Attitudes and Trends / Baobao Zhang	Baobao Zhang University of Oxford—Center for the Governance of AI; Yale University, Department of Political Science Baobao Zhang will present her report around the American public's attitudes toward artificial intelligence (AI) and AI governance, based on findings from a nationally representative survey of 2,000 American adults. As the study of the public opinion toward AI is relatively new, she and her team aimed for breadth over depth, with questions touching on: workplace automation, attitudes regarding international cooperation, the public's trust in various actors to develop and regulate AI, views about the importance and likely impact of different AI governance challenges, and historical and cross-national trends in public opinion regarding AI. Results provide preliminary insights into the character of US public opinion regarding AI. Zhang, Baobao and Dafoe, Allan, *Artificial Intelligence: American Attitudes and Trends* (January 9, 2019). Available at: https://isps.yale.edu/sites/default/files/files/Zhang_us_public_opinion_report_jan_2019.pdf.
November 7, 2019	How Neuroscience Can Help Computer Vision *and vice-versa	Summary: Artificial vision has often been described as one of the key remaining challenges to be solved before machines can act intelligently. Recent developments in a branch of machine learning known as deep learning have catalyzed impressive gains in computer vision—giving a sense that the problem of vision is getting closer to being solved. In this talk, I will provide a brief overview of recent deep learning developments followed by a critical assessment of our actual progress toward achieving human-level visual intelligence. I will discuss the implications of the successes and limitations of modern computer vision algorithms for biological vision and the prospect for neuroscience to inform the design of future artificial vision systems. Speaker bio: Thomas Serre is Associate Professor in Cognitive Linguistic & Psychological Sciences at Brown University. He received a PhD in Neuroscience from MIT in 2006 and an MSc in EECS from Télécom Bretagne (France) in 2000. Dr. Serre is Faculty Director of the Center for Computation and Visualization and Associate Director of Brown's animal behavioral core and the "SmartPlayroom."

November 7, 2019	How Neuroscience Can Help Computer Vision *and vice-versa	Thomas Serre bio (continued): Dr. Serre has served as an area chair for machine learning and computer vision conferences including CVPR, AAAI, and NeurIPS. He is currently serving as a domain expert for IARPA's Machine Intelligence from Cortical Networks (MICrONS) program and as a scientific advisor for Vium, Inc. He is the recipient of an NSF Early Career award as well as DARPA's Young Faculty Award and Director's Award. His research seeks to understand the neural computations supporting visual perception and has been featured in the BBC series "Visions from the Future" and appeared in several news articles (*The Economist*, *New Scientist*, *Scientific American*, *IEEE Computing in Science and Technology*, *Technology Review*, and *Slashdot*).
November 21, 2019	AI in libraries: case studies from Finland	Libraries all over the world are increasingly starting to apply artificial intelligence in their work. AI technologies, most notably machine learning, are used to support different aspects of library work, including creating metadata, enhancing logistical operations, and supporting information discovery. Join the RI-AI Meetup and Pirjo Kangas, Information Specialist from Humak University, Finland and Fulbright Finland Grantee 2019, to discuss recent developments with AI in Finnish libraries. Pirjo is currently a Fulbrighter at the University of Maryland libraries. Her current research is focused on exploring AI initiatives in Nordic and US libraries.
December 5, 2019	AI and Big Data: The Promise and Perils to Diversity and Fairness	Summary: It is no longer possible to have a career as a private citizen in the United States, and indeed much of the world, without the direct or indirect influence of complex algorithms churning over big data. As of today in the United States and most other countries, you have almost no rights to view your data, let alone ensure that your data are correct, nor do you have any right to inspect the algorithms which ultimately have large effects on your day-to-day personal and professional life. These algorithms and systems are demonstrably biased based on race, gender, ethnicity, religion, socioeconomic status, and sexual orientation for the few for which we can test. We will discuss the state of research on a number of subjects concerning algorithmic fairness and some of the existing and potential consequences and remedies.

December 5, 2019	AI and Big Data: The Promise and Perils to Diversity and Fairness	Speaker bio: Gabriele Fariello (former interim/transitional CIO here at URI) is an internationally recognized leader in building, improving, and turning around computational science, data science, and information technology organizations. He created and teaches the introductory survey course in machine learning and artificial intelligence at Harvard University. More here: https://people.fas.harvard.edu/~fariello/biosketch/.
February 3, 2020	Apply Machine Learning With Limited Real Data	Dr. Matt Wei is an assistant professor of oceanography at the University of Rhode Island (URI). He got his PhD from UC San Diego. His group uses remote sensing/ geophysical data and numerical models to study earthquakes and monitor nuclear tests. In this talk, Dr. Matt Wei would like to share his one-year experience of using machine learning in an area with limited real data. Machine learning is quite powerful but it has many imitations, which he recognized recently. By giving this talk, he hopes to get feedback from the community and inspire more machine learning projects at URI.
February 11, 2020	Harnessing Biomedical Data in a Quest to Understand Complex Genetic Disorders	At the beginning of the 21st century, we are experiencing the tremendous societal and economic impact of common diseases that are molecularly and genetically complex. These complex diseases include cancer, neurological disorders, chronic depression, heart disease, diabetes, and many others. Recent advances in the Next Generation Sequencing (NGS) technology have provided us with large volumes of data, revealing that many complex diseases are linked to the variations in the key genetic mechanisms, as compared to the data from healthy individuals. In this talk, I will introduce our recent work on understanding the effects of molecular mechanisms associated with complex genetic disorders, with the focus on studying how disease-associated changes can impact large-scale molecular networks and tissues. I will describe our recent projects where new machine learning methods were developed to decipher transcriptional signatures of the cell, discover novel mechanisms behind diabetes, and provide the first steps toward early diagnostics of chronic depression and suicidality.

February 11, 2020	Harnessing Biomedical Data in a Quest to Understand Complex Genetic Disorders	Dr. Dmitry Korkin is an Associate Professor and Director of the Bioinformatics and Computational Biology Program at Worcester Polytechnic Institute (WPI) in August 2014. Before coming to WPI, he was an Associate Professor at the University of Missouri-Columbia and the core faculty of Informatics Institute. Dr. Korkin did his postdoctoral research at the University of California San Francisco and Rockefeller University. He received his PhD in 2003 at the University of New Brunswick, Canada, and bachelor and masters at the Moscow State University, Russia. Dr. Korkin is a recipient of the NSF CAREER Award and the University of Missouri Junior Engineering Research Faculty of the Year award. He is a Senior Member of the International Society for Computational Biology (ISCB). His research is interdisciplinary and spans the fields of bioinformatics of complex disease, computational genomics, systems biology, and biomedical data analytics.
February 26, 2020	A Framework for Analyzing Spatial Networks for Utilities	Analyzing spatial networks for utilities (i.e., utility networks) such as electric, gas, or water networks has several critical societal applications and provides tremendous business value. For example, analysis may answer questions about the current state of the network (e.g., what valves need to be closed to fix a gas leak while minimizing the number of affected customers?), help to design future facilities (e.g., how many houses are fed by a transformer and can the transformer supply another house without overloading its capacity?), and help to organize business practices (e.g., create circuit maps for work crews to facilitate damage assessment after an ice storm). Analyzing utility networks is a challenging problem due to (1) the size of the data, which could have many tens of millions of network elements per utility, and billions of elements at the nationwide or continental scale, (2) modeling and analyzing utility assets at high fidelity (level of detail), and (3) the different analysis requirements across utility domains (e.g., water, wastewater, sewer, district heating, gas, electric, fiber, and telecom). This talk describes a framework for utility network analysis called the trace framework that has been implemented in ArcGIS Pro and ArcGIS Enterprise (10.6 and later). The trace framework features algorithms in a services-based architecture for addressing analysis tasks across a wide array of utility domains. It leverages a network model designed for utility networks. Unlike previous approaches that have focused on solving specific problems in specific domains, the trace framework provides a more general, scalable solution.

February 26, 2020	A Framework for Analyzing Spatial Networks for Utilities	Bio: Dev Oliver is a senior software development engineer at Esri, where he leads development efforts for the trace framework, a subsystem used for the analysis of utility networks (e.g., electric, water, gas) and trace networks (e.g., hydrography); the trace framework has been implemented in ArcGIS Enterprise and ArcGIS Pro. Dev graduated with his PhD in Computer Science from the University of Minnesota (2014) in the broad area of Spatial Computing. He also holds a master's degree in Computer Engineering from the University of Florida (2008) and a bachelor's degree in Computer Science from Macalester College (2004). His research and development interests are at the intersection of GIS and Computer Science (e.g., Spatial Networks, Big Data, Spatial Data Mining, Spatial Databases, and Spatial Data Summarization).
March 2, 2020	Mapping With Lidar To Guide Utilities And First Responders	Abstract : Light detection and ranging (LiDAR) technology has greatly advanced the capabilities of remote sensing to gather information on the 3D structures of objects and landscapes. LiDAR data have a wide variety of applications including modeling forest structure, general land cover mapping, and infrastructure mapping and assessment. Automated or semi-automated techniques are typically needed to extract information from LiDAR data across large areas; however, the immense size of these datasets makes them challenging to process. In this presentation, I will discuss projects in which we are using LiDAR for a variety of purposes including the development of models to assess infrastructure vulnerability to damage from trees, mapping utility infrastructure to improve risk assessments, and mapping building interiors to support public safety operations. For each project, I will focus on challenges that we have encountered and the solutions that we have found for working with LiDAR data.

March 2, 2020	Mapping With Lidar To Guide Utilities And First Responders	Bio: Dr. Jason Parent is an assistant professor in the University of Rhode Island's Department of Natural Resources Science. His research uses remote sensing and geospatial technologies to address problems related to natural resources and the environment. Current projects include characterizing forest conditions using LiDAR and unmanned aerial systems, assessing vegetation risk to infrastructure, and evaluating the effectiveness of utility vegetation management strategies. Dr. Parent has a PhD in Remote Sensing and Geospatial Science (2014) and a Masters in Earth Resources Information Systems (2006) from the University of Connecticut. Cosponsored by The URI Big Data Initiative (Library) URI College of Environment and Life Sciences.
April 24, 2020	From Kelp Forests to Coronavirus: an interactive webinar	OK. You've learned R in class. You use it once in a while to run an analysis or make a figure for a paper. Is that all? From kelp forests to coronavirus, Join Jarrett Byrnes, Associate Professor in Biology at UMass Boston, to talk about how his journey in Data Science has fundamentally changed the kinds of science he does. He'll also discuss some new directions in how he's using data science in R to build knowledge. Rather than just a tool for every now and again, he'll talk about how data science in R has become a useful part of his very existence. And, yes, this is an informal discussion/seminar for some good stories of how he learned to stop worrying and embrace R. And how you can, too! Speaker: Dr. Jarrett Byrnes, Assistant Professor of Biology—Marine Ecology, UMass Boston.
June 5, 2020	What today's AI adoption has led us to so far	Free Virtual Meetup: This talk will explore a few socio-economic and political questions that today's AI adoption has surfaced in the areas of production, governance, and labor management. Link will be provided later. Bohyun Kim, Chief Technology Officer and Associate Professor at University of Rhode Island Libraries.

June 11, 2020	Data Ethics related to COVID-19	RI-AI meetup is co-listing with Tech Collective RI. We will send the link 3 hours before. Data Ethics related to COVID-19 Tech Collective understands the power of collective collaboration and together, with our committees, has drafted a plan to support you. [Virtual Event] Our presenters are working on this agenda—Check back soon for details! About the presenter: Joan Peckman is a Professor of Computer Science and Campus-wide Coordinator of Big Data and Data Science Initiatives at the University of Rhode Island. Her research and teaching interests include databases, data modeling, Computer Science and Data Science education, diversity, and interdisciplinary (or convergent) engagement. She has earlier served as program director at the National Science Foundation (2008–2011), and chair of the Computer Science & Statistics Department (2011–2017). She led the development of the Data Science programs at URI. Doug Friedman is the Data Science Manager at Johnson & Johnson's Healthcare Technology Center in Providence, RI. There, he leads a team of data scientists solving some of the toughest analytical problems across the Johnson & Johnson enterprise—pharmaceutical, medical devices, commercial. He has a strong passion for open-source software and agile development. His work can be found at https://github.com/doug-friedman.

Endnotes

1. Lisa Federer, "Defining Data Librarianship: A Survey of Competencies, Skills, and Training," *Journal of the Medical Library Association* 106, no. 3 (July 2018): 296, https://doi.org/10.5195/jmla.2018.306; Franklin Sayre and Amy Riegelman, "Replicable Services," *College & Research Libraries* 80, no. 2 (March 2019): 267, https://doi.org/10.5860/crl.80.2.260.
2. Diane Geraci, Chuck Humphrey, and Jim Jacobs, *Data Basics: An Introductory Text*, preface iii (2012), http://3stages.org/class/2012/pdf/data_basics_2012.pdf.
3. Sarah Pugachev, "What Are 'The Carpentries' and What Are They Doing in the Library?," *portal: Libraries and the Academy* 19, no. 2 (2019): 209–14, https://doi.org/10.1353/pla.2019.0011.
4. Geraci, Humphrey, and Jacobs, *Data Basics*, 8.1.
5. Ibid., 9.8.
6. Samantha Guss, "A Studio Model for Academic Data Services," in *Databrarianship: The Academic Data Librarian in Theory and Practic*e, eds. Lynda M. Kellam and Kristi Thompson, (Chicago: ACRL, ALA 2016), 13.

7. John Burke, "Making Sense: Can Makerspaces Work in Academic Libraries?" (paper presented at ACRL 2015: Creating Sustainable Community, Portland, OR, March 25–28, 2015), 503, http://www.ala.org/acrl/acrl/conferences/acrl2015/papers.
8. Guss, "A Studio Model," 11–12.
9. Ibid., 21.
10. "Classes, Workshops, Training," NVIDIA Deep Learning Institute, last modified December 3, 2019, https://www.nvidia.com/en-us/deep-learning-ai/education/.
11. Peter Z. Vaillancourt, J. Eric Coulter, Richard Knepper, and Brandon Barker, "Self-Scaling Clusters and Reproducible Containers to Enable Scientific Computing," Cornell University, 2020, https://arxiv.org/abs/2006.14784.

Bibliography

Burke, John. "Making Sense: Can Makerspaces Work in Academic Libraries?" Paper presented at ACRL 2015: Creating Sustainable Community, Portland, OR, March 25–28, 2015. ACRL 2015 Proceedings: ACRL. http://www.ala.org/acrl/acrl/conferences/acrl2015/papers.

Federer, Lisa. "Defining Data Librarianship: A Survey of Competencies, Skills, and Training." *Journal of the Medical Library Association* 106, no. 3 (July 2018): 294–303. https://doi.org/10.5195/jmla.2018.306.

Geraci, Diane, Chuck Humphrey, and Jim Jacobs. *Data Basics: An Introductory Text*. 2012. http://3stages.org/class/2012/pdf/data_basics_2012.pdf.

Guss, Samantha. "A Studio Model for Academic Data Services." In *Databrarianship: The Academic Data Librarian in Theory and Practic*e, edited by Lynda M. Kellam and Kristi Thompson, 9–24. Chicago: Association of College and Research Libraries, a Division of the American Library Association, 2016.

NVIDIA. "Classes, Workshops, Training." NVIDIA Deep Learning Institute. Last modified December 3, 2019. https://www.nvidia.com/en-us/deep-learning-ai/education/.

Pugachev, Sarah. "What Are 'The Carpentries' and What Are They Doing in the Library?" *portal: Libraries and the Academy* 19, no. 2 (2019): 209–14. https://doi.org/10.1353/pla.2019.0011.

Sayre, Franklin, and Amy Riegelman. "Replicable Services for Reproducible Research: A Model for Academic Libraries." *College & Research Libraries* 80, no. 2 (March 2019): 260–72. https://doi.org/10.5860/crl.80.2.260.

Vaillancourt, Peter Z., J. Eric Coulter, Richard Knepper, and Brandon Barker. "Self-Scaling Clusters and Reproducible Containers to Enable Scientific Computing." Cornell University, 2020: 1–8. https://arxiv.org/abs/2006.14784.

Chapter 3

Artificial Intelligence, Machine Translation, and Academic Libraries:

Improving Machine Translation Literacy on Campus

Lynne Bowker, Maria Kalsatos, Amy Ruskin, and Jairo Buitrago Ciro

Introduction

One of the places on university campuses where artificial intelligence (AI) is appearing is in free online machine translation systems like Google Translate. International students are increasingly employing machine translation; however, they are not necessarily using it critically.[1] Can academic libraries help students develop the skills they need to be informed machine translation users? This chapter explores how AI-based machine translation systems work, considers why academic libraries are well placed to support international students with machine translation use, and presents the results of a pilot project to deliver a machine translation literacy workshop at two Canadian university libraries.

Machine Translation

Early machine translation systems tried to process language using bilingual dictionaries and grammar rules; however, this approach had limited success owing to the fact that language is inherently ambiguous and computers do not have the type of contextual knowledge that people use to interpret language.[2] For example, the French word *avocat* can be translated into English as either *lawyer* or *avocado*, depending on the context. Most English speakers can easily determine that contexts about food usually require the translation *avocado*, while those about court typically need the translation *lawyer*. However, people can also easily determine that the expression "that judge eats lawyers for breakfast" is correct, whereas a computer might be flummoxed.

In the 1990s, as it became easier to access large amounts of data, researchers moved away from linguistic approaches to machine translation and began to investigate data-driven approaches, such as statistical machine translation. Developers first fed the machine translation system with huge corpora of texts in language B to use as a language model. Next, they fed the system with enormous corpora of texts in language A along with their translations into language B (that had been translated by professionals) to use as a translation model. With reference to these training corpora, the systems tried to calculate the probability that a phrase in a new text should be translated the same way that it was translated previously.[3] Overall, statistical approaches to machine translation produced higher quality translations than did linguistic approaches.

Currently, neural machine translation is the dominant method. This approach also relies on enormous training corpora of previously translated texts, but it employs artificial neural networks, too. Neural machine translation incorporates artificial neural networks in the form of encoder-decoder frameworks with a source language attention model. One reason that neural machine translation outperforms statistical machine translation is that once the sentence from the source language has been processed by the encoder, the full context of the sentence is available to the decoder for consideration as to which target-language words and phrases should be suggested for the translation. Additional improvement was seen when a source-language attention model was added. Rather than accepting that all source-language words are equally important in suggesting all target-language words, the attention model demonstrates which source words are most relevant for hypothesizing target-language equivalents. Providing a more technical explanation of neural machine translation is beyond the scope of this chapter; however, accessible explanations of this AI-based technology exist in the literature.[4]

For this chapter, it is sufficient to note that neural machine translation systems produce higher quality translations than linguistic or statistical approaches. Previously, machine translations were of very poor quality and did not meet many translation needs. Improving those machine-translated texts was often more labor-intensive than translating them manually. Now, however, neural machine translation output offers a solid starting point that can be edited or possibly used unedited for some translation needs. It is not perfect and potential machine translation system users must develop good judgment and practical skills to know when and how to use this technology appropriately and effectively. In other words, users need to develop a new type of digital literacy: *machine translation literacy*.

Martin and Grudziecki describe digital literacy as "the awareness, attitude and ability of individuals to appropriately use digital tools and facilities to identify, access, manage, integrate, evaluate, analyse and synthesize digital resources, construct new knowledge, create media expressions, and communicate with others, in the context of specific life situations, in order to enable constructive social action; and to reflect upon this process."[5] This definition emphasizes critical thinking rather than technical competence. Like digital literacy, machine translation literacy is primarily a cognitive issue rather than a techno-procedural one. Using machine translation is easy; using it *critically* requires thought. In the context of online machine translation, important questions are *whether*, *when*, and *why* to use this technology. With regard to human-computer interaction, a key question is, "How can users interact with this tool to improve its output?" By asking such questions, people can become informed users of machine translation. The next section explores why academic libraries are well placed to help international students develop machine translation literacy skills.

Academic Libraries: Experts in Adapting to the Changing Needs of Their Communities

Campus units, such as an international office or a language-learning institute, often seem the best places to support international students. However, the library can also contribute to international students' academic success, including helping them to develop machine translation literacy skills. Academic libraries have experience in adapting their services to meet changing user demographics and in delivering other types of literacy training (e.g., information, media, digital literacy).

Academic Libraries and International Students

As communities evolve, libraries must monitor changes and adapt their services. University communities in Canada and elsewhere are experiencing increased internationalization. In 1997, Canada hosted fewer than one hundred thousand international students; however, by 2017, this number had increased five-fold.[6] Moving forward, Canada has announced a five-year strategy to expand and diversify academic internationalization.[7] Individual universities are also working to attract more international students. According to its strategic plan, "Destination 2020," the University of Ottawa seeks to double the number of international graduate students and increase the number of international undergraduates by 50 percent by 2020.[8] Concordia University reported that in 2018, international students comprised 20 percent of its student population, and over half of all students were not native English speakers.[9]

How are academic libraries addressing the needs of this linguistically and culturally diverse group? Click, Walker Wiley, and Houlihan conducted a systematic review of the literature on international students and academic libraries published between 1990 and 2014 and discovered that literature on this subject is burgeoning.[10] For instance, Bordonaro stresses that academic librarians need to begin by taking the time to understand how users from different countries view library resources, spaces, and services, and identifies intersections of library use and language learning, suggesting that learning to use library resources effectively can also lead to improvements in second language abilities.[11] Hughes et al. describe how four US and Australian universities are supporting international student transition into an unfamiliar academic and socio-cultural environment through actions such as providing wayfinding information in multiple languages, expanding the collection in additional languages, offering cultural awareness programs for library staff, offering targeted instructional support for international students, and organizing library activities during vacation periods since many international students remain on campus.[12] Meanwhile, Jackson and Sullivan present a dozen case studies detailing innovative projects to support international students, from creating research ambassadors to engaging with international students before they arrive on campus.[13]

Academic Libraries and Digital Literacy Instruction

Academic libraries are emerging as leaders in digital literacy and digital literacy instruction. For instance, Feerrar describes how the University Libraries at Virginia Tech created a new digital literacy framework to promote this concept on their campus.[14] Hallam, Thomas, and Beach at the University of Queensland Library, as well as Hensley and Bell at the University of Illinois Urbana-Champaign Library and Temple University Libraries respectively, describe the establishment of digital scholarship centres at these institutions and emphasize that digital scholarship considers the human, social, and cognitive dimensions alongside the technical one.[15]

It is not surprising that academic libraries are stepping into this role given that some degree of digital literacy is now required in every discipline. As a multidisciplinary unit, an academic library is well placed to offer digital literacy instruction across a university. Moreover, academic librarians have already developed instruction skills by delivering other types of literacy programs.

AI-based technologies are increasingly entering academic library life. In a recent research position paper from OCLC, Padilla observes that integrating AI with library research support and pedagogy presents multiple opportunities, noting specifically that the "potential afforded by technologies and methods is about enhancing the value of an existing service or presenting an opportunity to fill a gap."[16]

This edited volume explores AI in academic libraries, and the goal of this chapter is to investigate how one particular application of AI, machine translation, can be integrated into an academic library's offer of service in the form of machine translation literacy instruction. In their 2013 trend report, the International Federation of Library

Associations and Institutions predicted that machine translation would become a key trend in the global information environment and would consequently become part of the new digital paradigm in libraries.[17] The following sections describe a pilot project to deliver a machine translation literacy workshop for international students in academic libraries at two Canadian universities.

Machine Translation Literacy Instruction: A Pilot Project

The principal investigator and author (Bowker) is a professor at the University of Ottawa, cross-appointed between the School of Information Studies and the School of Translation and Interpretation. She was the 2019 Researcher-in-Residence at Concordia University Library, where she worked with research assistants and co-authors Kalsatos, Ruskin, and Buitrago Ciro to conduct the pilot project. Both Ruskin and Buitrago Ciro are international graduate students in information science, while Kalsatos relocated from the English-speaking province Ontario to the French-speaking province Quebec to pursue graduate studies in translation; as such, all have experience living and studying in a new linguistic and cultural environment.

Workshop Design

Getting support from key partners was critical to the workshop's success; therefore, the team adopted a community-based participatory research (CBPR) approach. CBPR is an approach that centers on researchers who work with a community to increase understanding of a given phenomenon and integrate the knowledge gained to improve the quality of life of community members.[18] Relevance and trust are key in CBPR. Communities are not enthusiastic about projects of limited interest or benefit to them and to which they have little input and limited access to the findings. Lack of trust and respect deter individuals from participating in research. However, if the research question addresses locally identified needs and if the design and methods actively engage community members, there will be greater trust and respect between researchers and communities.[19] In addition to having members with international student experience, the research team also invited the following units to collaborate on workshop content, to promote the events, and to provide facilities.

- Concordia University—Library, International Students Office, Continuing Education (unit responsible for instruction in English as a second language)
- University of Ottawa—Library, International Office, Academic Writing Help Centre

These partners contributed different knowledge and perspectives, and the result was the first draft of a collaboratively designed workshop intended to address international students' needs. The pilot workshop was forty-five minutes long with fifteen minutes for questions, and it focused on the six core elements outlined below. The workshop was presented in English using plain language.

1. Privacy/confidentiality. Information entered in a free online machine translation system does not disappear when the window closes. Instead, the machine translation company can keep and re-purpose the data. Sensitive information should not be entered into free online machine translation systems.

2. Academic integrity. When using machine translation for academic work, it is essential to cite and reference the original ideas that are borrowed, even if they are presented in another language. Furthermore, using machine translation for coursework may not be acceptable if it is contrary to the course learning objectives (e.g., language learning).

3. Potential for algorithmic bias. Machine translation systems learn from training corpora. If this material contains biases, then the tool may perpetuate them. For instance, there are reports that Google Translate generally skews toward masculine pronouns for words like *strong* or *doctor* and feminine pronouns for *beautiful* and *nurse*.[20]

4. Awareness of different tools. Google Translate is a well-known system, but there are others. Since each system is trained using different corpora, the various tools may not produce identical results. In particular, the language in question may affect a tool's performance. For instance, Yandex.Translate is strong for Russian, Baidu Translate works well with Chinese, and DeepL Translator could be a better choice for the more commonly used European languages (e.g., French, Spanish). Because machine translation systems require such large training corpora, these tools tend to produce higher quality results for widely used languages, but languages but may not work as well for languages that are less widely spoken. However, the systems are constantly learning, so their results may improve. It is recommended to try different systems and to not discount a system because of one poor performance.

5. Awareness of different translation tasks. Translation can be undertaken for numerous reasons. Machine translation is often helpful for understanding foreign-language texts, whereas it may not be suitable for producing texts for distribution unless the translation is first revised. Similarly, a machine translation system that has been trained using scientific texts may not translate legal documents well. It is important to consider purpose, content, and audience before deciding whether machine translation is a good option.

6. Improving the output by changing the input. Linguistic ambiguity is challenging for machine translation. One key way to improve machine translation output is to provide clear input. A translation-friendly writing style could include short sentences, the active voice, consistent terminology, no abbreviations, and no idiomatic expressions, humour, or culture-bound references.[21]

Pre-Pilot Test

The workshop was tested at Concordia with two international students who have different profiles to see if any elements needed significant modification. The demographic information of the students included

- one male and one female;
- one doctoral student and one undergraduate student;
- one sciences student (engineering) and one humanities student (business); and
- one Bangla speaker and one Mandarin speaker.

The students participated in a trial run of the workshop and provided feedback. Overall, the material was well received with both participants indicating that they had learned useful information. In terms of improvements, the students recommended reducing the specialized terminology and highlighting which machine translation tools were optimal for different languages.

Pilot

The workshop was piloted at the University of Ottawa and Concordia University. All fifty-six participants provided demographic information; however, only 73 percent of attendees completed a workshop evaluation. Table 3.1 summarizes the participants' profiles and table 3.2 highlights results from the workshop evaluations.

Table 3.1

Profile of workshop participants

	University of Ottawa	**Concordia University**
Date	October 2019	November 2019
Number of participants	27	29
Level	Undergraduate: 23 Graduate: 4	Undergraduate: 27 Graduate: 2
Discipline	Sciences (10), Humanities (17)	Sciences (6), Humanities (23)
Native languages	Mandarin (22), French (2), Arabic (1), German (1), Korean (1)	Mandarin (21), Farsi (2), Malay (1), French (1), Polish (1), Romanian (1)

Table 3.2
Responses to the end-of-workshop evaluation

Statements on workshop evaluation	**% at University of Ottawa who *agree* or *strongly agree* with the statement**	**% at Concordia University who *agree* or *strongly agree* with the statement**	**Average**
I learned new things about machine translation at this workshop.	76%	83%	79.5%
I feel confident that I can now use machine translation more effectively.	71%	73%	72%
I feel confident that I can improve machine translation quality using translation-friendly writing techniques learned at this workshop.	63%	66%	64.5%
I intend to increase my use of machine translation in my studies after this workshop.	78%	80%	79%
I will recommend this workshop to a friend or colleague.	80%	83%	81.5%
I would like to attend a more advanced follow-up workshop on machine translation.	41%	54%	47.5%

Discussion

As table 3.1 illustrates, the workshop attracted similar students at both institutions. Most participants were undergraduates, which could suggest that graduate students have already developed greater competence and confidence with regard to second-language writing, while undergraduates may need additional support such as translation tools.

More humanities students than sciences students participated at both institutions. This could be owing to the fact that humanities students have more writing assignments to complete and therefore have a greater need for writing support tools.

In both groups, most students were Mandarin speakers, with smaller numbers of other languages being represented. This corresponds to Canadian government data which indicate that most international students in Canada are from China.[22] Given that machine translation systems may produce different quality levels for different language pairs, it would be interesting to gather data pertaining to a wider range of languages to know whether machine translation presents a more promising option for some students than for others.

It appears that Concordia participants appreciated the workshop slightly more than did University of Ottawa participants given that a marginally higher number of Concordia respondents replied positively to each of the statements. The workshop was delivered for the first time at the University of Ottawa and was delivered several weeks later at Concordia. Between the two workshops, the authors conducted various seminars (e.g., for library staff), thus becoming more familiar with the material and delivery, which may explain why the Concordia participants responded more favourably.

Overall, it is gratifying to see that more than 80 percent of respondents would recommend the workshop to a friend, over 75 percent felt they learned something new about machine translation, and more than 70 percent felt more confident about using machine translation effectively. This increased level of knowledge and confidence could explain why nearly 80 percent of respondents report that they intend to use machine translation more often for future academic work and why less than half of the respondents feel a need to participate in a follow-up workshop.

There is room for improvement in helping students learn how to improve machine translation output by modifying the input. Less than two-thirds of the respondents felt confident about applying these techniques. One reason that the workshop may have been less successful in helping students to master translation-friendly writing is that it requires knowledge of the source language (e.g., Mandarin). This raises several points for consideration. It could be useful to customize workshops for different languages since effective translation-friendly writing techniques differ from one language to the next. Moreover, to maximize the workshop's effectiveness, it could help to partner with both foreign language and English as a second language teachers. Indeed, at Concordia University the authors were also invited to present on machine translation literacy to English-language teachers in Concordia's Continuing Education unit, and these teachers were enthusiastic about the potential of machine translation for meeting some of their students' needs. Partnerships between academic libraries and other campus units could help to better achieve the shared goal of setting students up for success.

Conclusion

According to Arlitsch and Newell,

> AI will transform library services, forever altering the mix of skills and tools needed to serve our users. At the same time, AI will change the lives of our users, and the dynamics of our communities. It is difficult to forecast exactly what, or how pervasive, these changes will be. But change is certain.[23]

In this chapter, we considered one AI application, machine translation, and how international students are using it. Arlitsch and Newell observe that as AI becomes increasingly embedded in our activities, it is important for academic libraries to prepare their own staff and to provide continuing education for their communities. The long-established positioning of academic libraries as a place for life-long learning represents an opportunity to be leveraged.[24] Machine translation literacy instruction is one example of this type of opportunity, and the authors hope that this chapter has advanced the conversation about the implications and applications of AI in academic libraries.

Endnotes

1. Paul Stapleton, "Google Translate may transform teaching of international students," Times Higher Education, August 27, 2019, https://www.timeshighereducation.com/opinion/google-translate-may-transform-teaching-international-students.
2. Andy Way, "Machine Translation: Where Are We at Today?" in *The Bloomsbury Companion to Language Industry Studies*, eds. Erik Angelone, Maureen Ehrensberger-Dow, and Gary Massey (London: Bloomsbury Academic, 2020), 313–14.
3. Way, "Machine Translation," 315–16.
4. Ibid., 316–17.
5. Allan Martin and Jan Grudziecki, "DigEuLit: Concepts and Tools for Digital Literacy Development," *Innovation in Teaching and Learning in Information and Computer Sciences* 5, no. 4 (2006): 255.
6. "International students by top 50 countries of citizenship," Immigration, Refugees & Citizenship Canada, accessed February 10, 2020, https://open.canada.ca/data/en/dataset/2bf9f856-20fe-4644-bf74-c8e45b3d94bd.
7. "Building on Success: International Education Strategy (2019–2024)," Government of Canada, accessed February 10, 2020, https://www.international.gc.ca/education/strategy-2019-2024-strategie.aspx?lang=eng.
8. "Destination 2020," University of Ottawa, accessed February 10, 2020, https://www.uottawa.ca/about/sites/www.uottawa.ca.about/files/destination-2020-strategic-plan.pdf.
9. "Fast Facts," Concordia University, accessed February 10, 2020, https://www.concordia.ca/about/fast-facts.html.
10. Amanda Click, Claire Walker Wiley, and Meggan Houlihan, "The Internationalization of the Academic Library: A Systematic Review of 25 Years of Literature on International Students," *College and Research Libraries* 78, no. 3 (2017): 328, https://doi.org/10.5860/crl.78.3.328.
11. Karen Bordonaro, *Internationalization and the North American University Library* (Lanham, MD: Scarecrow Press, 2013); Karen Bordonaro, *The Intersection of Library Learning and Second-Language Learning: Theory and Practice* (Lanham, MD: Rowman and Littlefield, 2014).
12. Hilary Hughes et al., "The role of the university library in supporting international student transition: Insights from an Australian-American case study," *Journal of Academic Librarianship* 44, no. 5 (2018): 582, https://doi.org/10.1016/j.acalib.2018.06.003.
13. Pamela Jackson and Patrick Sullivan, eds., *International Students and Academic Libraries: Initiatives for Success* (Chicago: Association of College and Research Libraries, 2011).
14. Julia Feerrar, "Development of a framework for digital literacy," *Reference Services Review* 47, no. 2 (2019): 91, http://dx.doi.org/10.1108/RSR-01-2019-0002.
15. Gillian Hallam, Amberyn Thomas, and Bill Beach, "Creating a Connected Future through Information and Digital Literacy: Strategic Directions at the University of Queensland Library," *Journal of the Australian Library and Information Association* 67, no. 1 (2018): 42, https://doi.org/10.1080/24750158.2018.1426365; Merinda Hensley and Steven Bell, "Digital Scholarship as a Learning Center in the Library: Building Relationships and Educational Initiatives," *College & Research Libraries News* 78, no. 3 (2017): 155, http://crln.acrl.org/index.php/crlnews/article/view/9638/11064.
16. Thomas Padilla, "Responsible Operations: Data Science, Machine Learning, and AI in Libraries" (OCLC Research Position Paper, Dublin, OH: OCLC Online Computer Library Center, Inc., 2019), 20, https://doi.org/10.25333/xk7z-9g97.
17. International Federation of Library Associations and Institutions (IFLA), *Riding the Waves or Caught in the Tide? Navigating the Evolving Information Environment. Insights from the IFLA Trend Report* (The Hague: IFLA, 2013): 4, accessed February 10, 2020, https://trends.ifla.org/insights-document.
18. Shauna MacKinnon, ed., *Practising Community-Based Participatory Research* (Vancouver: Purich Books, 2018).

19. Catherine Riffen et al., "Community-based participatory research: Understanding a promising approach to addressing knowledge gaps in palliative care," *Annals of Palliative Medicine* 5, no. 3 (2016): 219, http://apm.amegroups.com/article/view/10850/11576.
20. Johanna Monti, "Gender issues in machine translation: an unsolved problem?" in *The Routledge Handbook of Translation, Feminism and Gender*, ed. Luise von Flotow and Hala Kamal (London: Routledge, 2020).
21. Lynne Bowker and Jairo Buitrago Ciro, *Machine Translation and Global Research* (Bingley, UK: Emerald Publishing, 2019), 55–78.
22. "International students," Immigration, Refugees and Citizenship Canada.
23. Kenning Arlitsch and Bruce Newell, "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries," *Journal of Library Administration* 57, no. 7 (2017): 794, https://doi.org/10.1080/01930826.2017.1362912.
24. Arlitsch and Newell, "Thriving," 194.

Bibliography

Arlitsch, Kenning, and Bruce Newell. "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries." *Journal of Library Administration* 57, no. 7 (2017): 789–98. https://doi.org/10.1080/01930826.2017.1362912.

Arnold, Doug. "Why Translating is Difficult for Computers." In *Computers and Translation: A Translator's Guide*, edited by Harold Somers, 119–42. Amsterdam: John Benjamins, 2003.

Bordonaro, Karen. *Internationalization and the North American University Library*. Lanham, MD: Scarecrow Press, 2013.

———. *The Intersection of Library Learning and Second-Language Learning: Theory and Practice*. Lanham, MD: Rowman and Littlefield, 2014.

Bowker, Lynne, and Jairo Buitrago Ciro. *Machine Translation and Global Research*. Bingley, UK: Emerald, 2019.

Click, Amanda, Claire Walker Wiley, and Meggan Houlihan. "The Internationalization of the Academic Library: A Systematic Review of 25 Years of Literature on International Students." *College & Research Libraries* 78, no. 3 (2017): 328–58. https://doi.org/10.5860/crl.78.3.328.

Concordia University. "Fast Facts." Accessed February 10, 2020. https://www.concordia.ca/about/fast-facts.html.

Feerrar, Julia. "Development of a Framework for Digital Literacy." *Reference Services Review* 47, no. 2 (2019): 91–105. http://dx.doi.org/10.1108/RSR-01-2019-0002.

Government of Canada. "Building on Success: International Education Strategy (2019–2024)." Accessed February 10, 2020. https://www.international.gc.ca/education/strategy-2019-2024-strategie.aspx?lang=eng.

Hallam, Gillian, Amberyn Thomas, and Bill Beach. "Creating a Connected Future Through Information and Digital Literacy: Strategic Directions at the University of Queensland Library." *Journal of the Australian Library and Information Association* 67, no. 1 (2018): 42–54. https://doi.org/10.1080/24750158.2018.1426365.

Hensley, Merinda, and Steven Bell. "Digital Scholarship as a Learning Center in the Library: Building Relationships and Educational Initiatives." *College & Research Libraries News* 78, no. 3 (2017): 155–58. https://doi.org/10.5860/crln.78.3.9638.

Hughes, Hilary, Liz Cooper, Michael Flierl, Mary Somerville, and Niraj Chaudhary. "The Role of the University Library in Supporting International Student Transition: Insights from an Australian-American Case Study." *The Journal of Academic Librarianship* 44, no. 5 (2018): 582–94. https://doi.org/10.1016/j.acalib.2018.06.003.

Immigration, Refugees and Citizenship Canada. "International students by top 50 countries of citizenship." Accessed February 10, 2020. https://open.canada.ca/data/en/dataset/2bf9f856-20fe-4644-bf74-c8e45b3d94bd.

International Federation of Library Associations and Institutions (IFLA). "Riding the Waves or Caught in the Tide? Navigating the Evolving Information Environment. Insights from the IFLA Trend Report." Accessed February 10, 2020. https://trends.ifla.org/insights-document.

Jackson, Pamela, and Patrick Sullivan, eds. *International Students and Academic Libraries: Initiatives for Success*. Chicago: Association of College and Research Libraries, 2011.

MacKinnon, Shauna, ed. *Practising Community-Based Participatory Research*. Vancouver, BC: Purich Books, 2018.

Martin, Allan, and Jan Grudziecki. "DigEuLit: Concepts and Tools for Digital Literacy Development." *Innovation in Teaching and Learning in Information and Computer Sciences* 5, no. 4 (2006): 249-67. https://doi.org/10.11120/ital.2006.05040249.

Monti, Johanna. "Gender issues in machine translation: an unsolved problem?" In *The Routledge Handbook of Translation, Feminism and Gender*, edited by Luise von Flotow and Hala Kamal, 457-68. London: Routledge, 2020.

Padilla, Thomas. "Responsible Operations: Data Science, Machine Learning, and AI in Libraries." OCLC Research Position Paper, Dublin, OH: OCLC, Inc., 2019. Accessed February 19, 2020. https://doi.org/10.25333/xk7z-9g97.

Riffen, Catherine, Cara Kenien, Angela Ghesquiere, Ashley Dorime, Carolina Villaneuva, Daniel Gardner, Jean Callahan, Elizabeth Capezuti, and M. Carrington Reid. "Community-based participatory research: Understanding a promising approach to addressing knowledge gaps in palliative care." *Annals of Palliative Medicine* 5, no. 3 (2016): 218–24. http://apm.amegroups.com/article/view/10850/11576.

Stapleton, Paul. "Google Translate may transform teaching of international students." Times Higher Education, August 27, 2019. https://www.timeshighereducation.com/opinion/google-translate-may-transform-teaching-international-students.

University of Ottawa. "Destination 2020." Accessed February 10, 2020. https://www.uottawa.ca/about/sites/www.uottawa.ca.about/files/destination-2020-strategic-plan.pdf.

Way, Andy. "Machine Translation: Where Are We at Today?" In *The Bloomsbury Companion to Language Industry Studies*, edited by Erik Angelone, Maureen Ehrensberger-Dow, and Gary Massey, 311–32. London: Bloomsbury Academic, 2020.

Chapter 4

Incubating AI:

The Collaboratory at Ryerson University Library

Fangmin Wang, Aaron Tucker, and Jae Duk Seo

Introduction

According to a survey commissioned by the Canadian federal government, Canadians have reported feeling overwhelmed by the rate of technological change and are unsure about who they can trust given the rapid changes presented by artificial intelligence (AI) and machine learning (ML) technologies.[1] A study from the Royal Bank of Canada suggests there are significant opportunities for researchers and educators to focus their efforts on addressing the realities of AI and to help achieve the required public level of digital literacy to compete in a transformed global economy and society.[2] To this end, the Canadian Federation of Library Associations argues that libraries are particularly well suited to expand algorithmic literacy, initiate and sustain transparent and open-access infrastructures, and advocate for more informed and effective governmental policy.[3]

Discussing the innovative work of the University of Rhode Island Library's AI Lab, Bohyun Kim contends that libraries themselves must also focus on the future of AI as libraries' continued relevance depends upon the tasks that AI performs well, such as chatbots, cataloguing, and indexing.[4] However, based on the 2018 Ex Libris Future Library survey, a very small percentage of librarians ran AI-related operations in their libraries even though the interest in adopting AI is quite high.[5] As Arlitsch and Newell point out, libraries must leverage public trust and community need to provide life-long learning opportunities for a contemporary world that is increasingly interpenetrated by AI.[6]

Looking more closely at academic libraries, while Wheatley and Hervieux found that "81.5% of universities sampled have an identifiable research hub that focuses on artificial intelligence," they concluded that "very few university libraries collaborate with other units within their institutions on artificial intelligence initiatives" and that "startlingly few academic libraries have begun to engage in official projects or labs centered on artificial intelligence."[7] Learning from all this, the Ryerson University Library's Collaboratory has successfully established itself as a community-based facilitator, connector, and incubator of cross-disciplinary research where AI and other emerging research experimentation and collaboration can thrive. This chapter briefly details the Collaboratory's advent, the expansion of its AI-focused services, and its key organizational principles based on the Community of Practice (CoP) methodology. This chapter concludes with lessons learned that will hopefully inform other academic libraries looking to successfully implement their own AI research communities.

Opening the Collaboratory

Opened in the fall of 2017, the Ryerson Library Collaboratory is an incubator and space for supporting a wide range of collaborative and interdisciplinary research. Accessible to graduate students, sessional instructors, and tenured faculty, this research hub is home to 3D printers, soldering stations, smart sewing machines, dye sublimation printers, laser cutters, robotic arms, virtual reality technology, and more. The Collaboratory has enjoyed significant growth over the past two years: 215 members were accepted, and twenty-two research projects were hosted. The Collaboratory's thriving membership is made up of a highly diverse and multi-disciplinary group of researchers from more than forty-five different disciplines and programs.

As part of the library, it has been important from its very beginning that the Collaboratory be open and accessible to the entire researcher community at Ryerson. Since its inception, all faculty and graduate students at Ryerson were eligible to work in the space. If a researcher had a project in mind, they could send an email to the Collaboratory team or describe their project needs while filling out the membership registration form. An initial research support appointment would then be booked, and the researcher and their team could start meeting with a research project consultant from the Collaboratory.

Besides requesting support for research projects, there were several additional ways that researchers could engage with the Collaboratory. Each month, they could attend any of the numerous research talks and workshops that were held in the Collaboratory, such as the Wikipedia Edit-a-thon: Indigenizing Wikipedia workshop and the Introduction to Building an Immersive Environment in Virtual Reality workshop. Quite often, these research events would attract further membership; as a result, a number of researchers joined the Collaboratory in order to continue their engagement with the research community facilitated by the space. In addition, monthly orientation meetings with Collaboratory staff gave new members opportunities to further explore new research collaboration opportunities. This led to situations wherein the Collaboratory's strengths spread via word-of-mouth and many new members were convinced to join by existing members. Cultivating similar research interests among researchers through informal networking, this type of community engagement effort certainly helped the Collaboratory grow its community.

Although there is no research funding available through the Collaboratory at this point in 2021, the team has connected its members to funding opportunities available within Ryerson University, such as the library's Open Educational Resource Grant as well as the Learning and Teaching Grant from the Centre for Excellence in Learning and Teaching. In addition, the Collaboratory has also provided in-kind support via technology, resources, and expertise to a number of faculty members' research grant projects. Looking to the future, the Ryerson Library is at the planning stage to fund a Researcher-In-Residence program to support new faculty and emerging researchers for which the Collaboratory will provide essential research support and space.

Furthermore, in order to understand and meet the needs of the researcher community, a faculty advisory committee was created to provide advice on the development of this initiative. This proved to be an especially urgent concern for the non-computer science researchers who needed complex technical support in fields such as AI. To address this fact, current member and faculty member of the Department of Computer Science, Dr. Alex Ferwon, has played an important role as one of the faculty advisors of the Collaboratory, helping develop its technological capacity from the beginning. In addition, he has frequently connected his faculty colleagues and students with the Collaboratory. Although not every researcher in computer science or technology has been interested in collaborating on projects outside of their disciplines, the Collaboratory's focus on supporting interdisciplinary research collaboration attracted faculty and graduate students who wanted to explore partnership opportunities with humanities and social science researchers. To further facilitate such relationships, several graduate students in computer science and engineering were hired as Collaboratory research support consultants, including one of the co-authors of this chapter, Jae Seo. The students provided much-needed technical skills to non-computer science researchers in humanities and social science who were interested in exploring new technologies for their projects.

With this solid foundation in place, conversations began in 2019 about how AI technologies could be used to drive vibrant innovation, scholarship started to come to the forefront of the Collaboratory's activities, and the team realized that a coordinated strategy was needed to support the Ryerson community.

Targeting a Visiting Scholar in AI and Operating with the CoP Methodology

The Ryerson University Library first embarked on its investigation into artificial intelligence technology support and community-building for AI practitioners in early 2019. As an urban university in the heart of Canada's biggest economy—the City of Toronto—the Ryerson Library has an important role to play in supporting the innovation and use of AI technologies and the dissemination of knowledge about AI in society. The Library Information Technology Services (LITS) team has a history of adopting new technologies and cultivating innovation. In addition to supporting the library's programming and services, the LITS team's technology expertise and infrastructure are essential to the teaching, learning, and research of the Ryerson community.

There was a strong desire among students, instructors, and faculty across Ryerson to build a community to support AI and other interdisciplinary research activities. The team recognized that there was an opportunity for the Ryerson Library to lead the practice of AI knowledge dissemination and to facilitate research collaboration through its Collaboratory, acting as an organizer for AI conversations happening across the campus and beyond. As the head of the LITS team and lead of the Collaboratory, one author, Fangmin Wang, was inspired by the CoP methodology, which is based on organic, people-led growth and innovation. This approach has driven grassroots management and social innovation efforts in healthcare and social work.[8] Communities of Practice (CoP) are about informal, decentralized, people-led change enacted by strategies that include sharing circles, lunch and learns, training the trainer, and peer-led collegial knowledge transfer.

Figure 4.1
The basic layout of the Ryerson Library Collaboratory space. Photo by Ryerson Library Staff.

Wang began the initial planning for structuring the Collaboratory by conducting background research into key change agents in the AI community in Toronto and beyond and by reaching out to other leaders for informal discussions. Wang also attended targeted conferences to determine the best way forward. One of those first steps was connecting with Michael Ridley, librarian emeritus at the University of Guelph, who has decades of experience as a leader in academic libraries and computing. Ridley now dedicates his time toward his doctoral study of AI and machine learning at the Faculty of Information and Media Studies at Western University. After attending an AI talk led by Ridley in early 2019 at the Ontario Library Association Super Conference, Wang engaged Ridley in a conversation.[9]

Ridley saw the gap between the rapid development of AI technologies and the lack of understanding of AI in the Canadian public. He was passionate about the critical role that universities and libraries could play in disseminating knowledge involving AI to the public. Ridley and Wang recognized that it was imperative to create a community and that the Ryerson Library could become a leader in this mission through the experimentation and incubation of AI initiatives in its Collaboratory. With the support of Chief Librarian Carol Shepstone, Ridley was appointed as the Ryerson Library's Visiting Scholar to help develop its research capacity and strategic partnerships in the areas of AI and machine learning.[10] Ridley was provided a dedicated office within the Collaboratory in order to engage with the rest of the library team, the Ryerson faculty, and with students on various AI-related initiatives. The collaboration with Ridley attracted several Ryerson researchers to the Collaboratory; Aaron Tucker, one of the authors, was among the first to join.

Figure 4.2

Ryerson graduate student Nabila Abraham demonstrates CoPs as she gives an AI talk in the Ryerson Library Collaboratory. Photo by Ryerson Library Staff.

Using the CoP methodology and inspired by spaces such as the aforementioned Rhode Island University Library's AI Lab, five key principles have been used to guide the Ryerson Library Collaboratory's work in supporting AI technologies knowledge and capacity in partnership with various communities.

1. Seek out the connectors. As a first step, the Collaboratory prioritized working with early adopters and leaders in the space. This led to the appointment of Michael Ridley as the Ryerson Library's Visiting Scholar and the evolution of an AI Journal Club toward a more interdisciplinary approach. Under Ridley's leadership, the Ryerson University Library also established partnerships with Toronto Public Library and the Canadian Federation of Library Associations to start a pan-Canadian AI literacy project.[11]

2. Let the researchers and graduate students lead. Following a services-based model, the Collaboratory worked with key researchers, like Tucker whose work is the subject of this chapter's first case study. His research was a guide in terms of the technology required to support his and other researchers' work. Additionally, the authors invited graduate student researchers to lead workshops, as demonstrated by Tanya Pobuda's workshop, "Creating an Artificial Intelligence (AI) Chatbot," in July 2020. Empowering researchers in more informal sessions allowed the research members of the Collaboratory to drive community engagement while also encouraging knowledge mobilization across different disciplines.

3. Take an interdisciplinary approach**.** Wang and his colleagues fostered an environment that pulled researchers from many different areas into the Collaboratory space to undertake a wide variety of research projects. This was exemplified by the ChinaTown board game led by interior design professor Linda Zhang, which utilized 3D printing, laser cutting, and drone photography, and contributed to Zhang's ongoing ChinaTown Heritage Preservation project.[12] When combined with the AI work by Tucker and Pobuda, as well as the wealth of other research simultaneously being undertaken in the Collaboratory, scholars engaged in the lab were able to interact with, learn from, and further collaborate in multidisciplinary ways.

4. Embracing failures**.** Valuing interdisciplinary work means supporting scholars who have a wide range of technological skills, including many non-experts who require complex collaborations. When discussing AI research more specifically, it is essential to know that progress is very often halting and frustrating; however, as this chapter's two case studies will illustrate, what might be seen as a failure from a computer science perspective may provide insight into a humanities context. The Collaboratory has therefore made it a priority to support the incubation of projects and provide the infrastructure to iterate through inevitable roadblocks.

5. Early wins are key**.** Early wins, such as the two case studies explained in depth in this chapter, are critical for the ongoing success of a new AI initiative. The Collaboratory team very quickly realized such wins were a means to entice other interdisciplinary research into the space while also acting as guideposts to future infrastructure plans related to AI. The incubation of Tucker's work, in particular, has led the Collaboratory into a forthcoming collaboration between

York University and Ontario College of Art and Design that focuses on critical uses of computer vision in public art.[13] Again, it is important to note that "wins" often look like failures: AI projects may encounter issues in terms of technology, in-house expertise, and/or a lack of time. However, even so-called failures contribute to communal knowledge and future work.

This Criminal Does Not Exist: A Case Study in Collaborating on AI from a Humanities Perspective

One of the first projects supported by the Ryerson Library Collaboratory was This Criminal Does Not Exist (TCDNE), led by Aaron Tucker, a sessional instructor in the English department. TCDNE disrupts the logic of massive data extraction and processing, utilizing the machine learning technique of a generative adversarial network (GAN) to uncover the inherent biases within law enforcement facial recognition technologies (FRT). The initial conception of TCDNE was driven by Tucker's desire to show the types of faces most present within problematic datasets without replicating and/or erasing those faces. Using machine learning (ML) techniques, the project creates synthetic faces that the GAN determines to be representative of the 1,173 faces in the Multiple Encounters Dataset.[14] The portraits generated are primarily of African American males, which speaks to the types of faces that are typically over-represented in mugshot datasets. TCDNE is in reaction to a wealth of research and reporting about the known inherent biases of FRT. As the Georgetown Law Center on Privacy & Technology's report, "The Perpetual Line-up," insists, "Face recognition may be least accurate for those it is most likely to affect: African Americans."[15] By harnessing AI and ML, TCDNE highlights the racial and gendered vulnerabilities within FRT and the urgent need to address such biases within the constructions and applications of such systems. At this stage, it is important to note that this research is in the prototype stage and not ready for wide public distribution. Following recommendations made by Eve Tuck and K. Wayne Yang, TCDNE will proceed only once Tucker has spoken and potentially collaborated with the stakeholders and communities most affected by this research, including local activist groups, the Ontario Law Commission, as well as faculty at Ryerson University.[16]

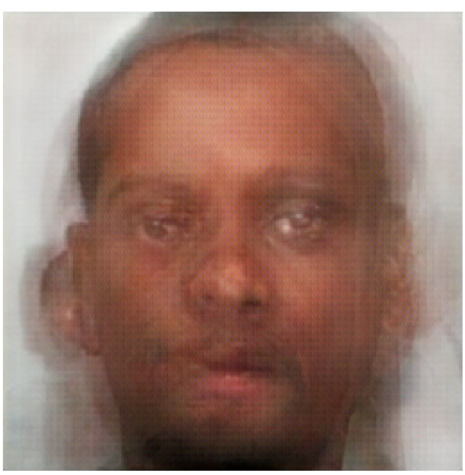


Figure 4.3

Example GAN-created image from This Criminal Does Not Exist.

The Collaboratory's support of TCDNE began with Tucker's collaboration with undergraduate student Kieran Ramnarine to build a face-generating GAN. The Collaboratory was essential in removing barriers that hindered early progress and in moving through the initial stages of the project's incubation. While the project began on Tucker's and Ramnarine's personal computers, Wang worked with Tucker to secure a machine-learning dedicated computer at the Collaboratory, providing the project with the much-needed infrastructure to undertake the weeks of uninterrupted machine learning that was required. The initial stages of the project at the Collaboratory involved detailed discussions about which hardware would best suit ML projects in general, effectively using TCDNE as a pilot project whose success would result in a scaling up of the Collaboratory's support of AI and ML.

Alongside this infrastructural support, the Collaboratory empowered TCDNE's lead researcher by providing Tucker with the opportunity to work closely with computer science graduate students. Tucker initially collaborated with Nabila Abraham to build basic FRT and learn the typical components of the software. Later, Tucker worked closely with Jae Seo to fine-tune earlier techniques learned from Abrahams, leading to the building of a complex suite of FRT and AI tools, which culminated in the Photogénie series.[17] The series explores the long history of celebrity faces within FRT databases, surfacing the white prototypicality that are inscribed into the technology by the faces used in its ML training.

Figure 4.4
Still from the Photogénie project. Aaron Tucker, 2020.

While TCDNE was ultimately successful, it was not without its challenges. As machine learning projects depend on long durations of computational processing, the project was repeatedly hampered by interruptions, ranging from the computer updating itself to power outages. Even when the GAN was able to continuously run, the results were uneven, mysteriously reverting to prior stages and rolling back weeks of progress. While some of these issues were eventually fixed with coding solutions and trial-and-error optimization, the most concrete solution was the upgrading of the computer hardware that greatly increased the speed at which ML could pace. Ultimately, however, the success of the GAN became secondary: while the faces generated as part of TCDNE were not as clear as a similar project within computer science might demand, the project's social justice arguments were still exemplified and perhaps enhanced by the painterly texture of the initial portraits. This was a key lesson for those in the Collaboratory: a humanities project that uses AI and ML requires patience and a tolerance for error and improvisation. The failures and imperfections may, in fact, generate more interesting and compelling research.

The Collaboratory's support of Tucker's work has been instrumental to its success. The granting of the computational hardware and technical expertise needed to undertake machine learning provided Tucker with the opportunity to fail quickly and nimbly on the way to success. Working together with Ramnarine, Seo, and Abrahams under the principles of CoP, Tucker gained technical knowledge while challenging the students' initial views of computer vision's objectivity and encouraging humanities-based paths into their research.

Moreover, Tucker's work has been instrumental to the Collaboratory's foundational thinking as it relates to expanding the support for future AI and ML projects across disciplines. Guided by CoPs, the Collaboratory's support of TCDNE showed that using machine learning to approach humanities projects can produce effective results and scholarship, even as a computer science perspective would likely dismiss the end results as a failure. TCDNE provided much of the template for the reciprocal educational and research environment that the Ryerson Library and its Collaboratory wished to create. By connecting Tucker to students and experts across disciplines, he activated his knowledge of AI and ML in iterative and effective ways.

Empowering Students and Communities within Interdisciplinary AI

The Collaboratory was able to further foster an interdisciplinary community by mentoring graduate students and researchers and by hosting events that are welcoming in nature. Through the mentoring between visiting scholar Ridley and Jae Seo, a graduate student in the Department of Computer Science at Ryerson University, graduate students were introduced to novel subjects outside their major while also gaining industry insights due to the practical experience of the mentors. Seo learned from Ridley how AI can be used in different organizations, including the Ryerson Library, and the extent to which those applications affect their respected institutions, such as in budget management.

The AI Journal Club is an initiative that enabled a variety of scholars to read and discuss scholarly articles related to AI. Created in the summer of 2018, it was primarily geared toward individuals who had computer engineering backgrounds. The first discussions tended to be very technical, and other non-engineering students and faculty were largely absent. In order to engage with the broader Ryerson community, Ridley and Seo shifted the sessions' focus to discuss the kind of change, positive or negative, that AI can bring to society. Following this, one of the Collaboratory's most successful talks was based on the paper, "Deep Fake and Cheap Fake."[18] Approximately thirty people attended the event, including students and faculty members from different departments. Participants raised a range of questions related to deep fakes, which, in turn, produced a lively discussion.

One of the less successful instances of the AI Journal Club was a session on "Green AI" which detailed how to make the whole ecosystem of machine learning more environment friendly.[19] While the general connection between environmental concerns and AI is extremely relevant, it didn't have the same instantaneous intrigue as the previous deep fakes talk and may have required different outreach methods. This failure has taught us valuable lessons on which topics related to AI have more appeal to the general public. Though it is certainly not ideal to only cover topics that have mass appeal, it is important to keep in mind the balance between the public's interest and the subject's urgency.

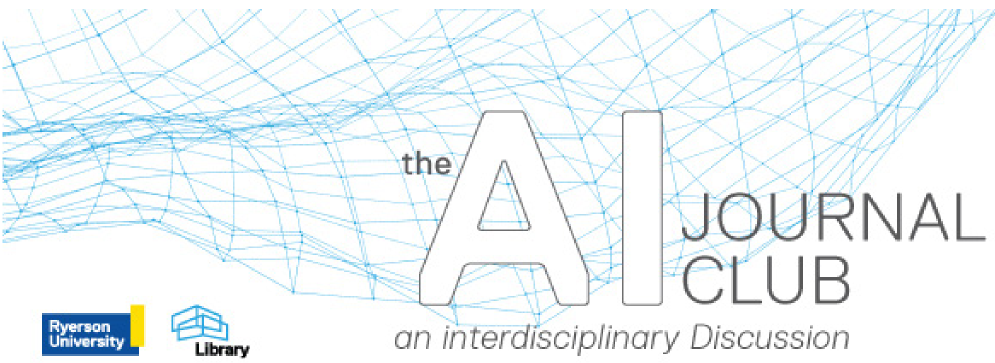


Figure 4.5
Logo for AI Journal Club. Design by Lee Chapman.

Building on the success of the AI Journal Club, Ridley and Seo, expanded the idea to another series of talks called "It's Not Artificial." This series, which began in October 2019, focused on real-world AI use-cases. One example from this series was "It's Not Artificial: Artificial Intelligence and You," where Melissa Hartwick, a former Strategy and Engagement Lead from Element AI, gave her opinions and thoughts on how AI will impact the general public.[20]

The Collaboratory goal of empowering graduate students does not end with providing them with opportunities to organize events. The Collaboratory is greatly served by the leadership of graduate students, such as Seo, who have become co-researchers who are instrumental in new AI initiatives. A concrete example of this would be the collaboration between Seo and Michael Carter-Arlt, a user experience researcher within the Collaboratory and a former master of digital media student who partnered with the Aga Khan Museum on the digital design and development of interactive exhibition content for the museum's Remastered Exhibition. The project leveraged web-based augmented reality (AR) technology to showcase the museum's upcoming exhibition. By taking charge of the technical aspects of the project, Seo learned how recent developments in ML and computer vision could be applied to the art industry. Ultimately, Seo acquired project management skills by learning how to effectively communicate with different stakeholders about the project's overall scope, while keeping the technical feasibility under consideration.

Overall, Seo's work at the Collaboratory is indicative of the rich and empowering experiences that graduate students receive by working in the space. The Collaboratory also gains a great deal from the leadership of graduate students like Seo. The collaborative efforts behind the AI Journal Club provided mentorships and learning experiences for Seo and knowledge mobilization for the Collaboratory. Seo's technical skills also greatly enhanced the work of other researchers, like Tucker. In turn, partners, like the Aga Khan Museum, are able to extend the service they provide to the general public by incorporating new and experimental technologies such as AR.

Discussion

The five key tenets listed prior were embodied in these case studies. The Ryerson Library, through the Collaboratory, *sought out the connectors*, looking for some of the key voices in the AI community, like that of visiting scholar Michael Ridley. The members of the Collaboratory purposely sought out AI-related events, listened to presentations, followed up with the key speakers, and found the key players in the local community and beyond. These early connectors helped build a community, and their generosity was a clear exemplification of CoPs at work. Because of these early connections, the Collaboratory was able to create a solid foundation for its burgeoning AI programs and services. It was through those conversations that the authors recognized the imperative of providing support for AI as an educational institution, as a library, and as members of Canadian society.

As shown by Tucker's work, the Collaboratory also *let the researchers lead the way.* Using a services-based model, the Collaboratory team worked hand-in-hand with other researchers to create the infrastructure, tools, and professional consultative support needed to make their research come to life. While creating a computing hardware system for running ML programs was new to the Collaboratory, the team was willing to experiment. Instead of trying to build something in anticipation of a need, they worked on a case-by-case basis to meet Ryerson researchers' requirements. In doing so, the Collaboratory created reproducible processes and resources that could benefit other researchers.

It wasn't enough for the Collaboratory to simply attract a handful of dedicated machine learning practitioners and computer scientists. Instead, Wang contemplated some of those early conversations with AI influencers, in turn fostering *an interdisciplinary space and initiative*. A common theme emerged: AI technologies should not be developed in a silo or an echo chamber. The Collaboratory best functions as an open and interdisciplinary space that provides a support system for the incubation of a range of projects. As both Tucker and Seo's work in the Collaboratory demonstrates, this support includes sessional instructors and graduate students who are empowered to experiment by the Collaboratory's resources.

The team also learned to not *worry about perfection.* This tenet was particularly key. How does a team looking to build a space like the Collaboratory start? The simplest answer: asking the scholars to identify the help they need. Leading the way can be an incredibly fruitful approach. Working with Tucker, as an example, made clear that the leads of projects know what they need.

The Collaboratory is just getting started, finding its way and engaging in experiments and trial and error. The team took pride in looking for the *early wins.* In one sense, that means taking on smaller ambitious projects that, even when they "fail," generate best practices and further communal knowledge for collaborative research spaces. In another sense, it means looking for those in your communities that are already doing innovative and publicly digestible work. Tucker's research is constantly at work in the Collaboratory as showcasing documentation from the ongoing stages of the project engages people and sparks fascinating discussions. In order to create a hub for activity, researchers must be willing to provide this kind of supportive sharing.

Conclusion

AI technologies are a key area for focus and attention among academic libraries. It is incumbent upon all academic librarians and staff to learn about AI technologies and to provide support grounded in an understanding of AI tools and best practices for their communities. Using the CoP methodology, the Ryerson Library, through the Collaboratory, embraced a strategy of gradual growth and an iterative approach to building a community. They went searching for the right people to inform their approach and connected with some of the leading voices from across the AI community. They were able to add value to these leaders by connecting them with students and faculty and by working alongside them to find and provide resources.

Emboldened by the Ryerson Library, the Collaboratory team built a set of key principles that could be used for future projects. They were committed to creating a diverse and interdisciplinary community. Ultimately, they realized community building is hard work and requires constant attention and commitment. They also came to understand that if they wanted everything to be perfect, they'd never take that first step. With some successful initiatives, they have developed a momentum that they hope will continue to expand AI research. Their community is fused together with a passion for this emerging technology and guided by an openness to experimentation. Their early AI pilot programs and initiatives were operationalized for wider use in the Ryerson Library and will hopefully provide a useful set of experiences for other libraries to follow.

Endnotes

1. "Canada's AI Imperative: From Predictions to Prosperity," Canadian Federal Government, accessed May 23, 2020, https://www.canada175.ca/en/reports/ai-imperative.
2. "Automation to Impact at Least 50% of Canadian Jobs in the Next Decade: RBC Research," RBC, accessed May 23, 2020, http://www.rbc.com/newsroom/news/2018/20180326-future-skills-rpt.html.

3. “Artificial Intelligence and Intellectual Freedom, Key Policy Concerns for Canadian Libraries,” Canadian Federation of Library Associations, May 2, 2018, accessed September 12, 2020, http://cfla-fcab.ca/wp-content/uploads/2018/07/CFLA-FCAB-2018-National-Forum-Paper-final.pdf.
4. Bohyun Kim, “AI and Creating the First Multidisciplinary AI Lab,” *Library Technology Reports* 55, no. 1 (2019): 16–22.
5. “Artificial Intelligence in the Library: Advantages, Challenges and Tradition,” Ex Libris, March 2018, accessed September 12, 2020, https://cdn2.hubspot.net/hubfs/2909474/Ex%20Libris%20Artificial%20Intelligence%20White%20Paper.pdf. The authors also considered Bruce Massis, “Artificial Intelligence Arrives in the Library,” *Information and Learning Science* 119, no. 7/8 (2018): 456–59.
6. Kenning Arlitsch and Bruce Newell, “Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries,” *Journal of Library Administration* 57, no. 7 (2017): 789–98.
7. Amanda Wheatley and Sandy Hervieux, “Artificial intelligence in academic libraries: An environmental scan,” *Information Services & Use* 39 (2019): 351–53.
8. Hubert Saint-Onge, *Leveraging Communities of Practice for Strategic Advantage* (New York: Routledge, 2017).
9. Michael Ridley and D. Pawlick-Potts, “Algorithmic Literacy: A Workshop for the Committed, the Curious, and the Confused (i.e. Everyone),” Ontario Library Association Super Conference, January 30–February 2, 2019.
10. “Introducing Ryerson University Library Visiting Scholar, Michael Ridley,” Ryerson Library, April 1, 2019, accessed September 16, 2020, https://library.ryerson.ca/blog/2019/04/introducing-ryerson-university-library-visiting-scholar-michael-ridley/.
11. The initial work for the in-progress initiative “AI for All” can be found at https://aiforall.ca/, accessed Sept 12, 2020. The project is currently undergoing a major overhaul and, as such, public details are minimal.
12. Michelle Grady, “Chinatown exhibition asks community what future heritage could look like,” *Ryerson University Today*, April 23, 2020, accessed September 12, 2020, https://www.ryerson.ca/news-events/news/2020/04/chinatown-exhibition-asks-community-what-future-heritage-could-look-like/.
13. The forthcoming collaboration between the Ryerson University Library Collaboratory, York University, and Ontario College of Art and Design critically explores AI and computer vision as it relates to issues of identity and representation; through a series of interlocking projects, the collaboration will generate public art that engages its audience in reflecting on their own data and AI practices.
14. “Special Database 32—Multiple Encounter Dataset (MEDS),” National Institute of Standards and Technology, May 10, 2010, accessed September 12, 2020, https://www.nist.gov/itl/iad/image-group/special-database-32-multiple-encounter-dataset-meds.
15. Clare Garvie, Alvaro M. Bedoya, and Jonathan Frankle, “The Perpetual Line-Up,” Georgetown Law Centre on Privacy and Technology, May 2016, accessed September 12, 2020, https://www.theperpetualineup.org.
16. Eve Tuck and K. Wayne Yang, “Unbecoming Claims: Pedagogies of Refusal in Qualitative Research,” *Qualitative Inquiry* 20 no. 6 (2014): 811–18; Eve Tuck and K. Wayne Yang, “Decolonization is not a metaphor,” *Decolonization: Indigeneity, Education and Society.* 1 no. 1 (2012): 1–40.
17. At the time of this writing, a longer version of the Photogénie series has been submitted to various festivals. However, an initial prototype of the project can be found at: “Photogénie: img_004_1.jpg [Prototype]” YouTube, last modified April 30, 2020, accessed 30 April 2020, https://www.youtube.com/watch?v=UECvBtcpihQ&t=19s. Original film still taken from *Some Like it Hot*. Dir. Billy Wilder. United Artists, 1959.
18. Britt Paris and Joan Donovan, “Deepfakes and Cheap Fakes,” Data & Society, September 18, 2019, accessed September 12, 2020, https://datasociety.net/library/deepfakes-and-cheap-fakes/.
19. Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni, “Green AI,” arXiv.org, Cornell University, August 13, 2019, accessed September 20, 2020, https://arxiv.org/abs/1907.10597.
20. “Global AI Software Provider,” Element AI, accessed August 24, 2020, https://www.elementai.com/.

Bibliography

Arlitsch, Kenning, and Bruce Newell. "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries." *Journal of Library Administration* 57, no. 7 (2017): 789–98.

Canadian Federal Government. "Canada's AI Imperative: From Predictions to Prosperity." Accessed May 23, 2020. https://www.canada175.ca/en/reports/ai-imperative.

Canadian Federation of Library Associations. "Artificial Intelligence and Intellectual Freedom, Key Policy Concerns for Canadian Libraries." May 2, 2018. Accessed September 12, 2020. http://cfla-fcab.ca/wp-content/uploads/2018/07/CFLA-FCAB-2018-National-Forum-Paper-final.pdf.

Element AI. "Global AI Software Provider." Accessed August 24, 2020. https://www.elementai.com/.

Ex Libris. "Artificial Intelligence in the Library: Advantages, Challenges and Tradition." March 2018. Accessed September 12, 2020. https://cdn2.hubspot.net/hubfs/2909474/Ex%20Libris%20Artificial%20Intelligence%20White%20Paper.pdf.

Garvie, Clare, Alvaro M. Bedoya, and Jonathan Frankle. "The Perpetual Line-Up." Georgetown Law Centre on Privacy and Technology. May 2016. Accessed September 12, 2020. https//:www.theperpetualineup.org.

Grady, Michelle. "Chinatown exhibition asks community what future heritage could look like." Ryerson University Today. April 23, 2020. Accessed September 12, 2020. https://www.ryerson.ca/news-events/news/2020/04/chinatown-exhibition-asks-community-what-future-heritage-could-look-like/.

Kim, Bohyun. "AI and Creating the First Multidisciplinary AI Lab." *Library Technology Reports* 55, no. 1 (2019): 16–22.

National Institute of Standards and Technology. "Special Database 32—Multiple Encounter Dataset (MEDS)." May 10, 2010. Accessed September 12, 2020. https://www.nist.gov/itl/iad/image-group/special-database-32-multiple-encounter-dataset-meds.

Paris, Britt, and Joan Donovan. "Deepfakes and Cheap Fakes." Data & Society. September 18, 2019. Accessed September 12, 2020. https://datasociety.net/library/deepfakes-and-cheap-fakes/.

Ridley, Michael, and D. Pawlick-Potts. "Algorithmic Literacy: A Workshop for the Committed, the Curious, and the Confused (i.e. Everyone)." Ontario Library Association Super Conference. January 30–February 2, 2019.

Royal Bank of Canada. "Automation to Impact at Least 50% of Canadian Jobs in the Next Decade: RBC Research." Accessed May 23, 2020. http://www.rbc.com/newsroom/news/2018/20180326-future-skills-rpt.html.

Saint-Onge, Hubert. *Leveraging Communities of Practice for Strategic Advantage.* New York: Routledge, 2017.

Schwartz, Roy, Jesse Dodge, Noah A. Smith, and Oren Etzioni. "Green AI." arXiv.org. Cornell University. August 13, 2019. Accessed September 20, 2020. https://arxiv.org/abs/1907.10597.

Tuck, Eve, and K. Wayne Yang. "Decolonization is not a metaphor." *Decolonization: Indigeneity, Education and Society* 1 no. 1 (2012): 1–40.

———. "Unbecoming Claims: Pedagogies of Refusal in Qualitative Research." *Qualitative Inquiry* 20 no. 6 (2014): 811–18.

Wheatley, Amanda, and Sandy Hervieux. "Artificial intelligence in academic libraries: An environmental scan." *Information Services & Use* 39 (2019): 351–53.

Chapter 5

Separating Artificial Intelligence from Science Fiction:

Creating an Academic Library Workshop Series on AI Literacy

Amanda Wheatley and Sandy Hervieux

Introduction

As artificial intelligence (AI) becomes more prominent in everyday conversations, it is critical for library users to develop basic knowledge and understanding of those technologies. In academic libraries, subject liaisons play a central role in the teaching of concepts like information literacy, open access, and research data management, in addition to their subject responsibilities. It should be no different, then, for these liaisons to take on the charge of learning and communicating the ways in which AI applications can change the research landscape for users. Two obvious challenges present themselves here: a lack of computational knowledge regarding AI and a lack of funds needed to develop a program of support. In this chapter, the authors present a model of AI engagement with users that is accessible for all knowledge backgrounds and is low-cost. With these two requirements in mind, a workshop series, Keeping Up with Artificial Intelligence, was created at the McGill University Library by two liaison librarians. The workshop series features three sessions: AI Literacy, AI Ethics and Bias, and AI in Research. The central premise of each workshop is to connect users from the community in discussions about AI so that all participants may leave with a richer understanding of the topic and how it may influence academic research.

Information Literacy and an AI Equivalent

The concept of information literacy (IL) has been constantly evolving since its initial invocation in the National Commission of Libraries and Information Science in 1974 by Paul Zurkowski.[1] The need to define information literacy and promote it as a skill needed by researchers was brought on by the rapidly increasing volumes of information being published. The amount of information available to the public predicated a desire on behalf of librarianship to assist in the navigation and evaluation of this information in order to support a more literate community.

In 1989, the American Library Association released the Presidential Committee on Information Literacy: Final Report, which posed the Information Age as one of the greatest challenges of the day.[2] They defined the information-literate person as one who "must be able to recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information."[3] For years, scholars have worked to expand on these preliminary remarks, leading to the current *Framework for Information Literacy for Higher Education* created by the Association of College and Research Libraries.[4] The *Framework*'s six threshold concepts—authority is constructed and contextual, information creation as a process, information has value, research as inquiry, scholarship as conversation, and searching as strategic exploration—have become canonized in the work of academic librarians.[5] Institutions have been hiring information literacy librarians for some time, a position that would require these librarians to promote and teach IL principles and values. Among these librarians, it is notable that each can approach the concept of IL and its definitions in varying ways. In Celene Seymour's 2012 study of IL librarians' work experiences, the author found that the rapid advances in technology and instant access to information were shifting the IL landscape for many librarians.[6]

It is not just information literacy that has taken the forefront of librarianship, the past couple of decades have also led to a rise in other literacy concepts, such as media, data, financial, and digital literacy. Librarians have taken up the challenge of teaching and supporting skills in these areas to varying degrees. Media literacy has become closely tied to IL practices, and following the United States' 2016 presidential election and the rise of "fake news," librarians all over the world saw occasion to promote media literacy and awareness to their communities. All one needs to do is take a look at LIS conference themes or programs over the last four years to be sure that this resurgence in media literacy (whether it is outwardly labelled that or not) has had an unprecedented impact on the profession.

While media literacy may have made the most noise among librarians, data and financial literacy have also made a resounding impact. Though financial literacy education tends to appear more in public libraries, it does have its space in the academic arena. Data literacy, on the other hand, has been ever-increasing in prominence. Tibor Koltay established data literacy as "a specific skill set and knowledge base, which empowers individuals to transform data into information and into actionable knowledge by enabling them to access, interpret, critically assess, manage, and ethically use data."[7] Koltay argued that even the label of data literacy was important as its lexical relationship with information literacy invokes the same importance to library users.[8] By comparison, the term *artificial intelligence literacy* has received much less recognition. Its use has been limited to small circles of education or computer science fields in recent years but has yet to become an established concept within librarianship. The authors of this chapter put forward AI literacy as a necessary distinction among the other concepts, especially digital literacy.

In 2019, researchers David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn posed five "Big Ideas" in AI that could be used as guidelines to teach students in K-12 programs about AI. These initial competencies are as follows: (1) computers perceive the word using sensors, (2) agents maintain models/representations of the world and use them for reasoning, (3) computers can learn from data, (4) making agents interact comfortably with humans is a substantial challenge for AI developers, and (5) AI applications can impact society in both positive and negative ways.[9]

The first substantive use of AI literacy, however, came about in the work of Duri Long and Magerko in 2020. Long and Magerko performed an exploratory review of AI literature in order to build their own set of competencies for learners.[10] They, too, cited the historical propagation of literacy concepts as a background for establishing AI literacy. Notably, only one source consulted on digital or data literacy was from a library journal. They define AI literacy, however, "as a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace."[11] Their conceptual framework includes five general themes with seventeen specific competencies and fifteen design considerations divided among said themes.

Where digital literacy might encompass topics like the evaluation and use of digital platforms, AI literacy is concerned with the advancing technologies that run those platforms. An AI-literate person can not only use their smartphone to access and create content on a social media platform, but they can also understand that certain features on those platforms are being embedded with AI technologies and speak to what those different programs can do. This separate distinction of AI literacy is built around the education of the public to better understand AI terminology and concepts and is encouraging them to become active participants with AI.

Building the Workshop Series

Given the importance of providing library users with artificial intelligence literacy, the authors created a series of three workshops called Keeping Up with Artificial Intelligence. Each two-hour session focused on different aspects of AI: AI Literacy, AI Ethics and Bias, and AI in Research. The workshops were developed at the McGill University Library in Montréal, Québec, Canada; notably, Montréal has played a significant role in the AI community as one of Canada's most prominent research cities dedicated to the topic.[12] In creating the workshops, the authors were determined to take an approach that welcomed all levels of knowledge toward AI. Neither author had extensive computational knowledge of AI prior to the development of this project; however, both authors have committed time over the past few years to strengthen their understanding. Participation in the Elements of AI online learning course and in discourse with the Montreal AI Ethics Institute are just two of the ways the authors worked to increase their knowledge of AI.[*]

Both authors began developing the project as a way to enhance AI experiences and conversations within the library. As there was no budget to provide funds, the project outcomes needed to be achievable with little to no capital spent. Thus, the three-part series was created as the best way to begin offering AI support on campus. The workshops were initially offered in March 2020 and were open to all in the McGill University community. A disclaimer on the workshop description let participants know that no computational knowledge of AI was needed to participate. As a result of the COVID-19 pandemic, only the first workshop, AI Literacy, was offered before the university closure, and it was attended by thirty-six participants. The full series of workshops was offered again in October 2020 in an online format. The authors delivered the three workshops via Zoom and thirty participants attended the series. Participants were not required to attend all workshops; they could choose to attend only one or multiple. The small number of participants enabled dynamic conversations in each of the workshops which were well received by the attendees.

AI Literacy

The first workshop in the series was an introductory session into AI technologies and knowledge competencies. As definitions of AI are crucial to the understanding of the topic, the workshop covered an exploration of AI terminology, participation in a Turing test, an evaluation framework, and the analysis of case studies on the use of AI in public practices.

Family Tree

In AI education, a metaphor often used to show understanding of the different capabilities of these applications is to compare artificial intelligence to human intelligence. However, AI technologies are often more complex and can be represented in their own frameworks. To convey these relationships, the authors created a family tree model to create a network of AI terminology as seen in figure 5.1. Specifically, the use of the infamous Kardashian family was used to not only provide levity to the complexity of AI relationships but also to make parallels between extended families and interdisciplinary

* Elements of AI, see: https://www.elementsofai.com/; Montreal AI Ethics Institute, see: https://montrealethics.ai/.

fields.† The break between AI and machine learning is a perfect example of contention within the field regarding whether these two areas should be interrelated or distinct. The family tree metaphor affords the learner the opportunity to see how these relationships intersect and diverge.

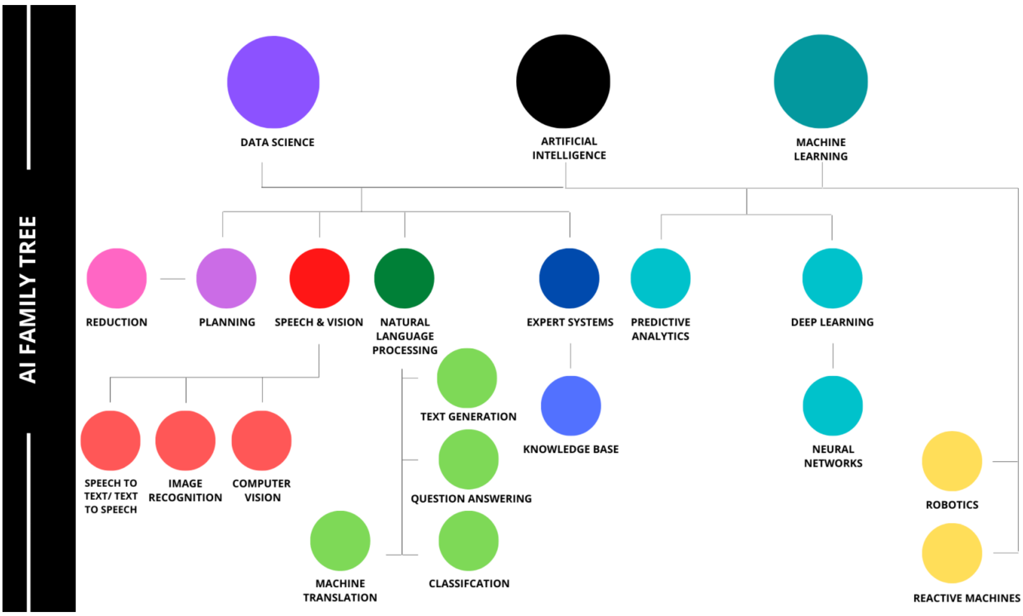


Figure 5.1

Family tree graphic of the relationship between artificial intelligence and other technologies.

ROBOT Test

Another important aspect of AI literacy is the ability to critically assess the information that is produced about AI technologies and the implications they can have. In order to help participants remember which aspects should be evaluated, an acronym was created. Much like the widely used CRAAP test, the ROBOT test enables its users to remember which questions to ask when encountering new information about AI technology.[13] ROBOT stands for reliability, objective, bias, ownership, and type. The authors created the ROBOT test to encourage its users to not only question and assess the information that they read about AI technologies but also to evaluate the technology itself. A complete outline of the ROBOT and the questions it prompts the user to ask can be found in table 5.1.

† The Kardashians are a well-known multi-generational blended family who were profiled in a reality television series, *Keeping Up with the Kardashians*, that ran for 20 seasons from 2007–2020.

Table 5.1
The ROBOT test components used as criteria when evaluating sources on AI

Reliability	• How reliable is the information about the AI technology? • If it's not produced by the party responsible for the AI, what are the author's credentials? Is there author bias? • If it is produced by the party responsible for the AI, how much information are they making available? Is information only partially available due to trade secrets? How biased is the information they produce?
Objective	• What is the goal or objective of the use of AI? • What is the goal of sharing information about it? To inform? To convince? To find financial support?
Bias	• What could create bias in the AI technology? • Are there ethical issues associated with this? • Are biases or ethical issues acknowledged? By the source of information? By the party responsible for the AI? By its users?
Ownership	• Who is the owner or developer of the AI technology? • Who is responsible for it? Is it a private company? The government? A think tank or research group? • Who has access to it? Who can use it?
Type	• Which subtype of AI is it? • Is the technology theoretical or applied? • What kind of information system does it rely on? • Does it rely on human intervention?

After introducing the ROBOT test to the participants, the authors distributed two recently published newspaper articles that discussed AI. They used the first article as an example to show participants how to properly assess the information and the technology. The participants were then asked to work in small groups to evaluate the second article according to the criteria of the ROBOT test.

AI Ethics and Bias

The second workshop in the series focused on the ethical issues and biases that can be present in AI. Given that participants were not required to attend all workshops, the librarians first provided an overview of AI terminology and the AI family tree to ensure that all attendees had the basic knowledge to be able to understand and participate in the workshop. They then introduced two newspaper articles as case studies to showcase possible ethical and privacy issues in AI. The first case study focused on an algorithm that was negatively biased toward people of colour.[14] The second one discussed privacy concerns with popular mass-market voice assistants, such as Google Home, Amazon Alexa, and Apple's Siri.[15] The authors provided the participants with some discussion questions about these case studies and encouraged them to discuss their implications for everyday life. The participants were then encouraged to brainstorm some possible solutions to ethical and privacy issues in AI technologies.

Following the introduction of ethical and privacy concerns, the authors provided an overview of some of the governmental legislation in place that can regulate the use of information and AI technologies. They also introduced participants to the OECD AI Policy Observatory.[16] The workshop participants worked in small groups to compare their own ideas to mitigate the ethical and privacy issues in AI with the legislation currently in place. The authors followed this activity with a presentation of the approaches of two different organizations engaged in AI: ElementAI and OpenAI. These two groups showed a contrast in producing transparent ethical statements regarding the development of their technology (the latter) and a commitment to the broad adoption of AI for economic benefit (the former).[17] Participants were encouraged to evaluate and reflect on the statements produced by these types of AI research groups and the products they developed. The authors wanted to stress that conversations on AI ethics should not just extend to concerns of how they are used (biased training data being a popular discussion example), they should also include considerations on the ethics of the researchers themselves.

AI in Research

The final workshop in the Keeping Up with Artificial Intelligence series, AI in Research, was not aimed at discussing research on AI but rather the implications of using AI applications during the research process. The authors asked participants to consider if they had ever used AI to augment their research process while discussing notable case studies. Some examples included the use of CrossRef technologies to review and reject journal submissions; TrendMD's article recommendation system, which is prominent in many journal databases; and the Semantic Scholar, which uses natural language processing for article searching.

Academic libraries are seeing a rise in the presence of digital scholarship on campuses and have grown to support this through the creation of librarian positions dedicated to the area and even the creation of technological hubs or centres where users can interact with new applications.[18] Already, the work that libraries have been doing is complementing the support of artificial intelligence as most applications involved in digital scholarship will include AI technologies. An example the authors explore in the workshop is that of Voyant, a text analysis and visualization tool geared toward the digital humanities.[*] Academics implementing Voyant in their research process may be unaware that it is a machine learning application that falls underneath the umbrella of AI.

While it is not feasible to ask a researcher to understand the code that makes up a program such as Voyant, it is important for them to understand the core roots of the software. Users of these programs are receiving aid to their research process; they may have a tacit understanding of the exchange of information that is taking place, but no focal knowledge of the subset of AI being utilized. The use of AI technologies is now prevalent

* Voyant, see: https://voyant-tools.org/.

in so many applications that it is nearly impossible to avoid its use in research. To help academics plan for AI augmentation, the authors prompted them to consider adding pointed questions on these topics into already existing research data management plans. A template of questions researchers can use in their planning process is presented below.

- How will you acknowledge the use of AI?
- Methods? Results? Discussion?
- What are the ethical and privacy concerns?
- If you are dealing with participant data, how will you protect, store, and anonymize it?
- Do you have permission to use this AI?
- How will you acknowledge it? Citations?
- Will your experiment and results be reproducible?
- Who is this AI available to?
- Open access versus proprietary?
- What is the level of oversight and verification?

Lessons Learned

The process of creating a workshop series proved to be a great learning experience for both authors. The participants' enthusiasm during the discussion portion of the March workshop was quickly noted; therefore, they decided to increase the length of the workshops from one-and-a-half hours to two hours to give attendees more time to engage in stimulating conversations. This change would have worked well for the in-person workshops; however, the authors noticed that conducting the session online sped up certain elements, like handing out papers and organizing participants into discussion groups.

The authors hope to increase marketing for the workshop series. After a high level of enthusiasm in the spring, attendance was lower for the fall sessions. It is possible that potential participants were tired of virtual meetings and classes; however, more efforts could be made to target groups with a possible interest in AI. The authors are investigating social media promotion through the library's accounts as well as advertising through student email lists.

While the workshops provide a basic understanding of AI literacy to its participants, the authors aim to build a more formal framework for artificial intelligence literacy. Similar to the ACRL *Framework for Information Literacy for Higher Education*, this framework would highlight the main competencies and attitudes that users should engage with. Special attention will be given to making the framework applicable and attainable by librarians in different institutions who have varying levels of knowledge of AI.

The authors will continue to offer the workshops at least twice a year and adapt the content to the ever-changing information landscape about artificial intelligence. The authors also aim to introduce a new workshop component where participants can engage directly with AI technologies, such as voice assistants, and evaluate their performance.

Conclusion

As society's interest and involvement in AI technologies continues to grow, the importance for individuals to be AI literate has never been higher. While academic librarians may feel unprepared or reluctant to teach their communities about artificial intelligence, it is possible to do so without expert knowledge or a computer science degree. Much like with digital, data, or media literacies, librarians can use their expertise and analytical skills to inform users and help them understand the implications of AI. Librarians are also known for their adaptability and willingness to learn, which makes them perfect candidates to adopt and teach AI literacy. The authors piloted a series of workshops that introduced users to the main topics related to AI, such as basic literacy, ethics and bias, and implications for research. While emphasis was placed on topics that would be relevant to the academic library community, similar workshops could be designed with different populations in mind. The authors intend to continue developing the Keeping Up with Artificial Intelligence workshop series and to build a framework for artificial intelligence literacy. They hope to continue bridging the divide between science fiction and reality.

Endnotes

1. Paul Zurkowski, *The Information Service Environment Relationships and Priorities. Related Paper No. 5.*, National Commission on Libraries and Information Science, 1974.
2. "Presidential Committee on Information Literacy: Final Report," American Library Association, 1989.
3. "Presidential Committee," American Library Association.
4. *Framework for Information Literacy for Higher Education*, Association of College and Research Libraries, American Library Association, February 9, 2015, accessed December 16, 2020, http://www.ala.org/acrl/standards/ilframework.
5. *Framework for Information Literacy*, ACRL.
6. Celene Seymour, "Ethnographic Study of Information Literacy Librarians' Work Experience: A Report from Two States," in *Transforming Information Literacy Programs: Intersecting Frontiers of Self, Library Culture, and Campus Community*, ed. Carrol W. Wilkinson and Courtney Bruch (Chicago: Association of College and Research Libraries, 2012), 55.
7. Tibor Koltay, "Data Literacy for Researchers and Data Librarians," *Journal of Librarianship and Information Science* 49, no. 1 (March 1, 2017): 10, https://doi.org/10.1177/0961000615616450.
8. Koltay, "Data Literacy," 11.
9. David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn, "Envisioning AI for K-12: What Should Every Child Know about AI?," *Proceedings of the AAAI Conference on Artificial Intelligence* 33, no. 01 (July 17, 2019): 9797–98, https://doi.org/10.1609/aaai.v33i01.33019795.
10. Duri Long and Brian Magerko, "What Is AI Literacy? Competencies and Design Considerations" (Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16, Honolulu, HI USA: ACM, 2020), https://doi.org/10.1145/3313831.3376727.
11. Long and Magerko, "What is AI Literacy?," 598.
12. Brendan Kelly, "Artificial Intelligence Expert Moves to Montreal Because It's an AI Hub," *Montreal Gazette* (September 5, 2020). https://montrealgazette.com/news/local-news/artificial-intelligence-expert-moves-to-montreal-because-its-an-ai-hub.
13. Sarah Blakeslee, "The CRAAP Test," *LOEX Quarterly* 31, no. 3 (October 1, 2004), https://commons.emich.edu/loexquarterly/vol31/iss3/4.
14. Ellora Thadaney Israni, "Opinion | When an Algorithm Helps Send You to Prison," *The New York Times* (October 26, 2017), sec. Opinion, https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html.

15. Kashmir Hill, “Activate This ‘Bracelet of Silence,’ and Alexa Can’t Eavesdrop,” *The New York Times* (February 14, 2020), sec. Technology, https://www.nytimes.com/2020/02/14/technology/alexa-jamming-bracelet-privacy-armor.html.
16. “The OECD Artificial Intelligence Policy Observatory,” accessed December 14, 2020, https://www.oecd.ai/.
17. “Responsibility,” ElementAI, accessed December 14, 2020, https://www.elementai.com/responsibility; “OpenAI Charter,” OpenAI, April 9, 2018, https://openai.com/charter/.
18. Alison Mackenzie and Lindsey Martin, *Developing Digital Scholarship: Emerging Practices in Academic Libraries* (London: Facet Publishing, 2016).

Bibliography

American Library Association. “Presidential Committee on Information Literacy: Final Report.” 1989.

Association of College and Research Libraries. *Framework for Information Literacy for Higher Education*. American Library Association. February 9, 2015. http://www.ala.org/acrl/standards/ilframework.

Blakeslee, Sarah. “The CRAAP Test.” *LOEX Quarterly* 31, no. 3 (October 1, 2004). https://commons.emich.edu/loexquarterly/vol31/iss3/4.

ElementAI. “Responsibility.” Accessed December 14, 2020. https://www.elementai.com/responsibility.

Hill, Kashmir. “Activate This ‘Bracelet of Silence,’ and Alexa Can’t Eavesdrop.” *The New York Times* (February 14, 2020). Sec. Technology. https://www.nytimes.com/2020/02/14/technology/alexa-jamming-bracelet-privacy-armor.html.

Israni, Ellora Thadaney. “Opinion | When an Algorithm Helps Send You to Prison (Published 2017).” *The New York Times* (October 26, 2017). Sec. Opinion. https://www.nytimes.com/2017/10/26/opinion/algorithm-compas-sentencing-bias.html.

Kelly, Brendan. “Artificial Intelligence Expert Moves to Montreal Because It’s an AI Hub.” *Montreal Gazette* (September 5, 2020). https://montrealgazette.com/news/local-news/artificial-intelligence-expert-moves-to-montreal-because-its-an-ai-hub.

Koltay, Tibor. “Data Literacy for Researchers and Data Librarians.” *Journal of Librarianship and Information Science* 49, no. 1 (March 1, 2017): 3–14. https://doi.org/10.1177/0961000615616450.

Long, Duri, and Brian Magerko. “What Is AI Literacy? Competencies and Design Considerations.” Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16. Honolulu, HI USA: ACM, 2020. https://doi.org/10.1145/3313831.3376727.

Mackenzie, Alison, and Lindsey Martin. *Developing Digital Scholarship: Emerging Practices in Academic Libraries*. London: Facet Publishing, 2016.

OECD.AI Policy Observatory. “The OECD Artificial Intelligence Policy Observatory.” Accessed December 14, 2020. https://www.oecd.ai/.

OpenAI. “OpenAI Charter.” April 9, 2018. https://openai.com/charter/.

Seymour, Celene. “Ethnographic Study of Information Literacy Librarians’ Work Experience: A Report from Two States.” In *Transforming Information Literacy Programs: Intersecting Frontiers of Self, Library Culture, and Campus Community*, edited by Carrol W. Wilkinson and Courtney Bruch, 45–71. Chicago: Association of College and Research Libraries, 2012.

Touretzky, David, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. “Envisioning AI for K-12: What Should Every Child Know about AI?” Proceedings of the AAAI Conference on Artificial Intelligence 33, no. 01 (July 17, 2019): 9795–99. https://doi.org/10.1609/aaai.v33i01.33019795.

Zurkowski, Paul. *The Information Service Environment Relationships and Priorities. Related Paper No. 5*. National Commission on Libraries and Information Science, 1974.

Chapter 6

Do Students Dream of Electric Cats (or Dogs)?:

Using Robotics for a Unique Exam Week Activity in the Library

Jonathan Scherger, Juliana Espinosa, Autumn Edwards, Chad Edwards, Bryan Abendschein, and Patricia Vander Meer

Introduction

Academic libraries regularly provide fun activities for students during exam weeks in an effort to reduce the stress that most students feel around the end of the term. Most of these activities involve some component of distraction, whether it be providing stress balls, board games, coloring books, or another diversion. A common offering is the opportunity to engage with a therapy animal, usually either a dog or a cat. Animals used in library activities are typically trained to deal with the public, have handlers that watch over their interactions with students, and are provided by an organization with a mission to provide pet therapy to the public.

Using animals to reduce stress and anxiety in humans has a long history in behavioral science. Animal-assisted therapy (AAT) has its roots in the work of Freud, Levinson, and as far back as Florence Nightingale in the 1800s.[1] Such therapies were linked to improvements in patients' cardiovascular, psychological, and cognitive health, as well as outcomes of treatment and hospitalization. There is a clinical distinction between AAT and Animal Visitation.[2] Visitation is a short-term intervention, whereas AAT is a longer-term, scheduled process with a trained therapist.[3] Visitation is more in line with how therapy animals have been used in libraries.

Activities with live animals began appearing during exam weeks in academic libraries between 2010 and 2011.[4] Outcomes of these activities frequently showed that therapy animals demonstrated the ability to reduce stress and anxiety in students. Multiple studies at universities and academic libraries seem to support the effectiveness of this approach.[5]

There are several concerns that relate to bringing live animals into a library, including shedding and defecation.[6] Although the positive effects of providing therapy animals often outweigh any negative factors, robotic animals, specifically dogs and cats, eliminate the biological concerns of handling a live animal and may offer a similar level of stress relief to students suffering from exam-related stress. A partnership between the University Libraries and faculty from the School of Communication, including co-directors of the Communication and Social Robotics Labs (COMBOTLABS) at Western Michigan University (WMU), examined that question using Ageless Innovation's Joy for All Companion Pets during a collaborative study conducted at Waldo Library during the fall 2019 semester final exam week. Western Michigan University is a Carnegie Higher Research Activity Doctoral University with a total enrollment of 21,470 students as of fall 2019.

Literature Review

The first historical instance of a robotic animal may have been a pigeon that moved by steam power created by Archytas of Tarentum during the third century BC.[7] While it would be a stretch to consider that pigeon a "pet," people have been creating artificial versions of companion animals for centuries. For example, a metallic robotic pet dog by the name of Sparko appeared in 1940 as a companion to a humanlike robot called Elektro, created by Westinghouse, which debuted at the 1939 World's Fair in New York.[8] The dog, which was powered by electric motors, was capable of walking, sitting up, and begging.

Bandi brought the Tamagotchi to the public in 1996.[9] The Tamagotchi did not physically look like a pet, but the software in the device was designed to simulate caring for a live animal. A number of robotic pets followed, including Tiger's Furby and Sony's AIBO.[10] Melson, Kahn, Beck, and Friedman synthesized data from three studies examining the effects of the robotic dog, AIBO, with human populations of different ages.[11] The results indicated that children as well as adults interacted with the robot in the same manner as a living animal.

Even though toys like AIBO move and behave like pets, they still visually look like robots. Ugobe's Pleo went a step further with a lifelike dinosaur toy that the company described as "autonomous life."[12] The lifelike nature of Pleo is critical to understanding how humans react to robotic animals, as Rosenthal-von der Pütten et al. demonstrated in their study examining emotional reactions to robots.[13] In the study, the researchers showed subjects a series of videos of someone interacting with the Pleo pet. One video showed someone in a friendly interaction with the robotic dinosaur, while a second showed the robot being abused. Study participants experienced negative feelings while viewing the abuse video, which suggests that it is possible for humans to have feelings for a robot they perceive as living.[14]

While the literature does not mention studies using robot pets as therapy for college students, robot pets have been commonly used in therapeutic applications. A study using the cat NeCoRo with dementia patients attempted to demonstrate that robotic pets could replace living pets in providing comfort through animated engagement instead of using a plush cat toy.[15] Another robot pet, the robotic harp seal PARO, was specifically designed for therapeutic uses.[16] In an Australian study, PARO, which can react to user movements via sensors, demonstrated the ability to improve perceptions of pleasure in dementia patients when compared to patients who participated in a reading group with other people.[17]

According to the Ageless Innovation's website, Hasbro introduced the Joy for All Companion Pets in 2015, first with a cat and then in a dog in 2016.[18] Similar to the PARO, the Joy for All robot pets use sensors to detect external movements and react with sounds and movements of their own. In his piece on the use of Joy for All robot pets at the Veterans Administration Palo Alto Health Care System (VAPAHCS) in Palo Alto, California, writer S. C. Stuart describes how some interviewed veterans ascribed feelings or associations from former pets onto the robotic dogs and cats.[19]

Several of the authors involved in this study participated in a previous project between the University Libraries and the COMBOTLABS in which usage of a telepresence robot (TR) on loan from the COMBOTLABS was piloted in the main library. The project consisted of several related studies: (1) COMBOTLABS and library student assistants invited patrons to interact with the robot, learn about the technology, and provide feedback; (2) use of the robot was tested by librarians for several public services applications; (3) perceptions were solicited from library staff and librarians regarding the usefulness of the technology and its applications in libraries before and after exposure to the robot and training in its operation. One of the project's findings was that "a TR can offer academic libraries a chance to showcase an emerging, engaging technology to its community."[20]

Planning

The authors met two months prior to the event to make decisions on dates, times, and location. The University Libraries offered to provide space, marketing support, and student staffing. COMBOTLABS provided the robotic pets, Ageless Innovation's Joy for All Companion Pets. Five cats and four dogs were obtained through funding provided by a Western Michigan University College of Arts and Sciences Discovery and Dissemination Award (CDDA). The School of Communication faculty took the lead on preparing a proposal and the participant consent form that the group submitted to the WMU Institutional Review Board (HSIRB). The Board granted approval as an expedited study given that the data was to be collected anonymously. Several meetings of two or more of the authors subsequently took place to address more detailed logistical issues and to address considerations that arose as the event days approached.

Librarians and the School of Communications faculty jointly created pre- and post-interaction surveys. The University Libraries provided supplies, such as hard copies of the surveys and HSIRB consent form, clipboards, and pens/pencils. The group staffing the events collected more than 100 paper surveys during the two nights the event was held at the library.

The authors created the coding for the survey prior to the first night and began the data entry process during the second night of the event. After the event, the authors finished entering the remaining data. The authors also recorded any informal observations about the event they had made while they were still fresh in their minds. COMBOTLABS student employees sorted through and input open-ended comments into an Excel spreadsheet. To ensure that everyone was able to access the various responses, all survey data was placed in a secure shared folder. The School of Communication faculty analyzed all recorded responses and reported the results at the conclusion.

A major advantage to the Communications faculty in working with the University Libraries on this project was the ability to utilize the University Libraries' marketing and social media options. The title and particulars regarding the event were given to the University Libraries' marketing team in order to create a campaign that would coincide with the advertising for the twenty-four-hour exam hours at the main library. Lead-time was important in order to advertise actively via social media, the University Libraries' calendar, the university's events calendar, and the campus newspaper. The marketing team created posters and flyers around the theme of a "petting zoo" of robotic animals. In line with the team's marketing strategy, the event was posted on Facebook and Instagram. Table tents and posters were also displayed throughout the main library a week prior to the event.

Figure 6.1
The animals on the promotional poster are appropriately portrayed in a robotic fashion, a la *The Day the Earth Stood Still*, with red eyes.

Methodology of the Study

The authors chose to hold the robot event between 8:00 p.m. and 10:00 p.m. during two consecutive nights, Sunday and Monday, of finals week. These nights fell toward the beginning of the libraries' twenty-four-hour Fall Finals schedule, where the main library typically stays open to students and staff until the end of exams. Based on an analysis of headcount and gate count statistics from previous finals weeks, Sunday and Monday seemed an opportune time to catch students looking for a break from studying. Evening hours are often high traffic in most academic libraries at that time of the semester, especially once regular classes have ended and just before scheduled exams begin.

A corner of the main library was chosen for the event location due to its high visibility and the ability for open interactions. Stanchions helped to designate the interaction space, as well as provide for crowd control in a mostly open area on the first floor of the library. The interaction space was staged with multiple tables, while comfortable lounge chairs and ottomans created a casual feel for the engaging interactions with the robot pets. Several chairs with desks were placed near the space for students to fill out the surveys. Photocopies of the HSIRB form, the research survey, clipboards, and pens were available for distribution. Counts of necessary materials had to be estimated as the University Libraries had not attempted a similar event previously.

Figure 6.2

Each robot has a unique pet name tag in order to personalize the units with typical dog and cat names, such as "Mittens," "Scout," "Patches," and "Bear."

Each of the authors staffed the event on both nights, along with two student employees from the main library's User Services department. Everyone was given brief training in welcoming participants and explaining the optional study. A critical key to the explanation was to avoid using terminology such as "stress reduction" with potential participants in order to prevent influencing the study results. (Similar language was also intentionally omitted from any materials or postings by the University Libraries' marketing team.)

While the event team all participated in various duties, the University Libraries faculty and students, and the Communications faculty, organically broke into two groups. University Libraries faculty and students welcomed participants, managed survey distribution and collection, and monitored the interaction space. Communications faculty took photos, engaged students in other parts of the library to encourage participation, and tallied data from the collected surveys. Student employees also maintained the interaction space, replacing batteries and re-arranging the pet robots after each interaction to ensure that pets looked available for the next group.

During the two nights the event was offered at the library, students were invited to interact with any of the nine battery-operated robots, which resembled and exhibited behavior like cats or dogs, including realistic heartbeat, purring and/or barking, and movement in response to touch and sound. Students were also invited to take part in the optional study consisting of informed consent, a pre-test prior to interaction, and a post-test at the conclusion of their visit. The surveys included a combination of closed-ended questions and open-ended prompts inquiring about participants' perceptions of the robot pets and their experiences interacting with them.

Results

Responses from students indicated the element of animal-like technology greatly enhanced the relaxation factor of the experience. Corresponding comments include, "I enjoyed the robots more than I thought I would. I really like [that] the cat purred and moved.… I felt like Biscuit and I had a special bond" and "I felt better about my finals after this event. I miss my dog at home now!"

The primary goal of this project was to reduce student stress during a challenging time of the semester. The results of the study indicated that a number of students appreciate library events designed to alleviate their stress. Students reported enjoying the opportunity to be "kids" again for a little while, something to keep in mind when planning activities. One student commented in the follow-up survey, "I was surprised by how much their interaction actually made me happy and excited. They responded the way I wanted them to and that was super fun." This study received more positive reviews than the previous study with the telepresence robot, which received mixed reviews from students.[21]

Despite mostly positive reactions, there were some mixed or negative comments, usually related to a sense of uncertainty about the robots' realism. "I have a puppy at my apartment, so this is rather close, but real animals would be better. It honestly kind of freaked me out" and "It was weird as I was very aware that it was not a real animal and did not find [I was] comforted or happy while petting them" were two of the comments that stood out as questioning or rejecting the robot animals as substitutes for their living counterparts.

A subset of students focused their interactions on the robot as opposed to the pet experience. Students were observed inspecting joints, testing various responses by the robotic pets (for example, waiving arms to see if it would trigger the dogs to bark), and feeling for wires and sensors. While the event was not intended as a showcase for robotics, the event attracted a few enthusiasts that were simply curious rather than interested in gaining any relaxation from the experience.

Conclusion

Offering an innovative relaxation activity with the robot pets proved to be a positive experience for both the attendees and the authors. Enlisting a department outside of the library allowed the authors to take advantage of different skills and knowledge in terms of technology, research practices, and experience with students when creating events. Utilizing robot pets, in particular, did attract students for a variety of reasons, including curiosity, connection to something that reminds them of their own pet, or the experience of participating in an empirically driven study.

Working with people across the campus involves a cycle of creating, implementing, and refining. For the project outlined in this chapter, specifically, the authors regularly met as a team to envision the event parameters, and then all worked to carry out a shared plan. When there is collective involvement in the planning and execution of an event, it can generate more ownership over the outcomes and lead to genuine, transferable enthusiasm. This type of collaboration inspires group members to continue working together to improve their original idea. For example, for the robot petting zoo outlined in this chapter, the authors continued to collaborate after the event and are working to implement several changes for the next iteration.

Enhancement to the study's design is being considered. In addition to gauging attendance as an indicator of success, the authors plan to utilize another factor that can be employed when an event is offered multiple times, such as rate of enthusiasm from an initial visit to a repeat visit. Progress can be built on this project's research element by collecting physical markers of stress through pulse oximeters during the pre/post-test surveys. A secondary consideration to repeating the event is whether the close contact with an inanimate object will create additional anxiety for students because of perceived risks of coronavirus infection, even after effective treatments are developed and available. A study comparing this study's results to a similar experiment with a robot that could easily be sanitized might yield data that would indicate whether perceptions of relaxation have now been altered by the perception of a robot's potential as a virus vector. The authors would also like to explore how altering the location of the event in the library may influence students' perceptions of the event and/or affect their reported benefits.

Endnotes

1. Lorraine Ernst, "Animal-Assisted Therapy: An Exploration of Its History, Healing Benefits, and How Skilled Nursing Facilities Can Set Up Programs," *Annals of Long-Term Care: Clinical Care and Aging* 22, no. 10 (October 2014), https://www.managedhealthcareconnect.com/articles/animal-assisted-therapy-exploration-its-history-healing-benefits-and-how-skilled-nursing.
2. "Animal-assisted interventions: Definitions," American Medical Veterinary Association, accessed May 22, 2020, https://www.avma.org/resources-tools/avma-policies/animal-assisted-interventions-definitions.
3. Katherine Connor and Julie Miller, "Animal-assisted Therapy: An In-depth Look," *Dimensions of Critical Care Nursing* 19, no. 3 (2000): 20.
4. Jo Ann Reynolds and Laurel Rabschutz, "Studying for Exams Just Got More Relaxing-Animal-Assisted Activities at the University of Connecticut Library," *College & Undergraduate Libraries* 18, no. 4 (2011): 360–62. https://doi.org/10.1080/10691316.2011.624934.
5. Mary Renck Jalongo and Theresa McDevitt, "Therapy Dogs in Academic Libraries: A Way to Foster Student Engagement and Mitigate Self-Reported Stress during Finals," *Public Services Quarterly* 11, no. 4 (February 2015): 264, https://doi.org/10.1080/15228959.2015.1084904; Amber Lannon and Pamela Harrison, "Take a Paws: Fostering Student Wellness with a Therapy Dog Program at Your University Library," *Public Services Quarterly* 11, no. 1 (2015): 21, https://doi.org/10.1080/15228959.2014.984264.
6. Reynolds and Rabschutz, "Studying for Exams," 365.
7. Kristen E. Schaefer et al., "The Future of Robotic Designs: Trends from the History of Media Representations," *Ergonomics in Design* 23, no. 1 (January 2015): 13, https://doi.org/10.1177%2F1064804614562214.
8. Clive Gifford, *A Robot World: Discovering Amazing Robots and their Robotic Powers* (Hauppauge, NY: Barron's Educational Services, 2018), 8.
9. Erika Wolf, "11 Fascinating Facts About Tamagotchi," Mental Floss, May 29, 2020, https://www.mentalfloss.com/article/624667/tamagotchi-facts.
10. Arielle Pardes, "The Second Coming of the Robot Pet," *Wired* (July 1, 2019), https://www.wired.com/story/the-second-coming-of-the-robot-pet/.
11. Gail F. Melson et al., "Robotic Pets in Human Lives: Implications for the Human-Animal Bond and for Human Relationships with Personified Technologies," *Journal of Social Issues* 65, no. 3 (2009): 562, https://doi.org/10.1111/j.1540-4560.2009.01613.x.
12. Keith Shaw, "Pleo: The Robot You'll Love," *Network World* 24, no. 49 (2007): 18.
13. Astrid M. Rosenthal-von der Pütten et al., "An Experimental Study on Emotional Reactions Towards a Robot," *International Journal of Social Robotics* 5, no. 1 (2012): 30.
14. Rosenthal-von der Pütten, "An Experimental Study," 29.
15. Alexander Libin and Jiska Cohen-Mansfield, "Therapeutic Robocat for Nursing Home Residents with Dementia: Preliminary Inquiry," *American Journal of Alzheimer's Disease and Other Dementias* 19, no. 2 (March/April 2004): 111–12, https://doi.org/10.1177/153331750401900209.
16. Kazuyoshi Wada and Takanori Shibata, "Living with Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House," *IEEE Transactions on Robotics* 23, no. 5 (2007): 973, https://doi.org/10.1109/TRO.2007.906261.
17. Wendy Moyle et al., "Exploring the Effect of Companion Robots on Emotional Expression in Older Adults with Dementia: A Pilot Randomized Controlled Trial," *Journal of Gerontological Nursing* 39, no. 5 (May 2013): 49, https://doi.org/10.3928/00989134-20130313-03.
18. "Ageless Innovation: Our Story," Ageless Innovation, accessed August 22, 2020, https://joyforall.com/pages/our-story.
19. S. C. Stuart, "How (Furry) Robot Companions Help Aging Veterans," PCMag.com, August 17, 2018, https://www.pcmag.com/news/how-furry-robot-companions-help-aging-veterans.
20. LuMarie Guth and Patricia Vander Meer, "Telepresence Robotics in an Academic Library: A Study of Exposure and Adaptation among Patrons and Employees," *Library Hi Tech* 35, no.3 (2017): 418.
21. Guth and Vander Meer, "Telepresence Robotics," 413.

Bibliography

Ageless Innovation. "Ageless Innovation: Our Story." Accessed August 22, 2020. https://joyforall.com/pages/our-story.

American Medical Veterinary Association. "Animal-assisted interventions: Definitions." Accessed May 22, 2010. https://www.avma.org/resources-tools/avma-policies/animal-assisted-interventions-definitions.

Connor, Katherine, and Julie Miller. "Animal-assisted Therapy: An In-depth Look." *Dimensions of Critical Care Nursing* 19, no. 3 (2000): 20–26.

Ernst, Lorraine. "Animal-Assisted Therapy: An Exploration of Its History, Healing Benefits, and How Skilled Nursing Facilities Can Set Up Programs." *Annals of Long-Term Care: Clinical Care and Aging* 22, no. 10 (October 2014): 27–32. https://www.managedhealthcareconnect.com/articles/animal-assisted-therapy-exploration-its-history-healing-benefits-and-how-skilled-nursing.

Gifford, Clive. *A Robot World: Discovering Amazing Robots and their Robotic Powers.* Hauppauge, NY: Barron's Educational Services, 2018.

Guth, LuMarie, and Patricia Vander Meer. "Telepresence Robotics in an Academic Library: A Study of Exposure and Adaptation Among Patrons and Employees." *Library Hi Tech* 35, no.3 (2017): 408–20.

Jalongo, Mary Renck, and Theresa McDevitt. "Therapy Dogs in Academic Libraries: A Way to Foster Student Engagement and Mitigate Self-Reported Stress during Finals." *Public Services Quarterly* 11, no. 4 (February 2015): 254–69. https://doi.org/10.1080/15228959.2015.1084904

Lannon, Amber, and Pamela Harrison. "Take a Paws: Fostering Student Wellness with a Therapy Dog Program at Your University Library." *Public Services Quarterly* 11, no. 1 (2015): 13–22. https://doi.org/10.1080/15228959.2014.984264.

Libin, Alexander, and Jiska Cohen-Mansfield. "Therapeutic Robocat for Nursing Home Residents with Dementia: Preliminary Inquiry." *American Journal of Alzheimer's Disease and Other Dementias* 19, no. 2 (March/April 2004): 111–16. https://doi.org/10.1177/153331750401900209.

Melson, Gail F., Peter H. Kahn, Jr., Alan Beck, and Batya Friedman. "Robotic Pets in Human Lives: Implications for the Human-Animal Bond and for Human Relationships with Personified Technologies." *Journal of Social Issues* 65, no. 3 (2009): 545–67. https://doi.org/10.1111/j.1540-4560.2009.01613.x.

Moyle, Wendy, Marie Cooke, Elizabeth Beattie, Cindy Jones, Barbara Klein, Glenda Cook, and Chrystal Gray. "Exploring the Effect of Companion Robots on Emotional Expression in Older Adults with Dementia: A Pilot Randomized Controlled Trial." *Journal of Gerontological Nursing* 39, no. 5 (May 2013): 46–53. https://doi.org/10.3928/00989134-20130313-03.

Pardes, Arielle. "The Second Coming of the Robot Pet." *Wired* (July 1, 2019). https://www.wired.com/story/the-second-coming-of-the-robot-pet/.

Reynolds, Jo Ann, and Laurel Rabschutz. "Studying for Exams Just Got More Relaxing-Animal-Assisted Activities at the University of Connecticut Library." *College & Undergraduate Libraries* 18, no. 4 (2011): 359–67. https://doi.org/10.1080/10691316.2011.624934.

Rosenthal-von der Pütten, Astrid M., Nicole C. Krämer, Laura Hoffmann, Sabrina Sobieraj, and Sabrina C. Eimler. "An Experimental Study on Emotional Reactions Towards a Robot." *International Journal of Social Robotics* 5, no. 1 (2012): 17–34. https://doi.org/10.1007/s12369-012-0173-8.

Schaefer, Kristen E., Jeffrey K. Adams, Jacquelyn G. Cook, Angela Bardwell-Owens, and Peter A. Hancock. "The Future of Robotic Designs: Trends from the History of Media Representations." *Ergonomics in Design* 23, no. 1 (January 2015): 13–19. https://doi.org/10.1177%2F1064804614562214.

Shaw, Keith. "Pleo: The Robot You'll Love." *Network World* 24, no. 49 (December 2007): 18.

Stuart, S. C. "How (Furry) Robot Companions Help Aging Veterans." PCMag.com. August 17, 2018. https://www.pcmag.com/news/how-furry-robot-companions-help-aging-veterans.

Wada, Kazuyoshi, and Takanori Shibata. "Living With Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House." *IEEE Transactions on Robotics* 23, no. 5 (2007): 972–80. https://doi.org/10.1109/TRO.2007.906261.

Wolf, Erika. "11 Fascinating Facts About Tamagotchi." Mental Floss. May 29, 2020. https://www.mentalfloss.com/article/624667/tamagotchi-facts.

Part II
Collections and Discovery

Chapter 7

Subjectivity and Discoverability:

An Exploration with Images

Catherine Nicole Coleman, Claudia Engel, and Hilary Thorsen

Introduction

What is in an image? The answers, of course, depend on who you ask. For many reasons, generating descriptions of images is challenging, not the least because any documentation of the image is in and of itself an artifact of its time.

In order for an image collection to be accessible, it is essential to communicate about its content. Cultural heritage institutions as custodians of image collections have to navigate the difficult line between the impossibility of a "perfect" image description and the charge of making their collections available for a public audience. A collection that is not catalogued is neither easily accessible nor discoverable. The responsibility falls onto the cataloguers and metadata librarians, trained experts in creating bibliographic descriptions.

Traditional cataloguing methods are expensive, time-intensive, and require specialized knowledge, particularly image cataloguing, which requires specific skills.[1] Libraries with large and unique collections commonly face a significant backlog of cataloging work that continues to grow with time. These two factors drive the need for libraries to find means to assist or augment the work of the cataloguer. Recent experiments suggest that artificial intelligence (AI) systems and computationally trained classifiers may be poised to automate much of that work or change our approach to information discovery entirely.

The potential of AI techniques to enrich, extend, and influence how and why we describe objects in the library seems particularly significant for images. Computer vision has been at the center of research into artificial intelligence because it supports the identification of objects and patterns without resorting to a text-based description. Why should libraries do without a tool that has proven so successful in commercial search engines?

This interdisciplinary study places traditional methods of cataloguing alongside experiments in AI. It explores how an image is understood in different contexts, what kind of image description is relevant to different audiences, and how images are described for purposes of preservation, information exchange, discovery, and research. To begin to understand how artificial intelligence can support cataloging work and benefit researchers, one needs to understand what cataloguers do, what researchers are looking for, and what classification and description mean in the context of AI.

Our relationship to technical objects, including those using machine learning (ML) techniques, is almost magical: we ask them to perform a task without knowing how that task is performed. This study looks at the affordances of technology to support and augment image discovery after first exploring the processes and purposes of image classification in the library in comparison with the approach used by researchers working with a large collection of archaeological images. The comparison reveals the many decision points throughout the process of describing an image collection: from the point of data gathering to formal cataloging and preservation to image retrieval, discovery, and, ultimately, reinterpretation through the research process.

The Challenge of a Formal Image Description

Choices about how collections are made discoverable have profound consequences. They cascade from the metadata specialist to the discovery environment and from there to the researcher who visits the online catalog. They affect the needs and practices of discovery, collection development, and technical services. Making these choices, however, raises a series of complex challenges.

For the cataloguer and the user, one of the main limitations of image description is the translation of images from their visual to textual form to be discoverable in the library catalog. Cataloguers need to anticipate how a user might search for an image, which can be problematic for a number of reasons. Library description standards may not capture domain-specific terms that a researcher might use. Resources devoted to describing a collection and the scale of the collection may not permit item-level description, thus impeding discovery.

Furthermore, finding and sharing images across institutions is significantly more difficult than finding and sharing books. Images are described differently by libraries, archives, and museums. Libraries typically use standards geared toward bibliographic materials. Archives often rely on description at a collection and item level using standard cataloging fields and descriptive fields. Museums tend to use local description standards and extensive curatorial research. Because images can be cataloged so differently, cross-institutional search is difficult. This deficit has long been recognized, making the possibility of more standardized cataloging methods appealing.[2]

Even within the library itself, the discovery and retrieval of images are hampered by cataloguing models that are based on the form and structure of publications. Cataloguing creates metadata for resources according to certain rules. The metadata records of today rely on authors, titles, and subjects of works as the primary access points for materials based on the card catalog created by Charles Ammi Cutter in the mid-nineteenth century. While, more recently, cataloging standards like Resource Description and Access have attempted to expand cataloging practices beyond books, they still rely on elements provided in published materials, including title, publisher, and date to supply ways of accessing the materials and ensure that the catalogue records created can be reused by other libraries.[3] These elements may not be readily attainable in images. Since there is often no direct association between the title and the work, subject access—the more general category-based query rather than the search for an individual item—is often standard for image retrieval.

Though item-level description could improve image discovery significantly, it is often not possible.[4] The most granular level of image description provides titles, names of creators, subjects, dates of creation, location, formats, size, quantity and extent, relationships with other people or organizations, and contextual information about the purpose of the image. While metadata may be supplied, item-level descriptions can require the cataloguer to examine images individually and possibly even perform additional research to provide more contextual information. As a result, subject access is often supplied based on the collection description rather than the item. Assigning subjects at the collection level may mean that images may only be retrieved as a group and single image retrieval may not be possible.

Because the quality and amount of metadata provided with images varies widely, cataloging images has been called a form of *translation* by Note.[5] According to Note, elements and concepts "perceived" by the cataloguer in the image are transformed into a controlled vocabulary. Whereas published materials commonly have structured metadata integrated into their form, such as publisher and publication date, image retrieval depends on how well the cataloguer has captured the image textually. Even though the vocabulary itself is controlled, the mapping of terms to the object is unavoidably subjective. It also is malleable, as terms are often corrected or revised over time to respond to the needs of users and researchers.[6] Nonetheless, subjects described by the cataloguers remain the primary entry points into a photographic collection.

A Case Study

Çatalhöyük, a nine thousand-year-old neolithic site in central Turkey, recognized as World Heritage by UNESCO, is possibly one of the oldest urban settlements in human history.[7] Archaeologists at Çatalhöyük pioneered a reflexive approach to archaeological practice, known as post-processual archaeology, in which information is permanently open to re-interpretation by both scholars and the public.[8] The reflexive archaeological method not only acknowledges the subjectivity of the archaeological researcher but also of the archaeological processes (decisions made during the excavation based on both, subsequent assumptions and new questions arising from continuous interpretation) and it strives to bring transparency to them.

During the twenty-five years of the project, the excavation was extensively documented to have a record of the processes and interpretations. This resulted in almost five terabytes of digital material, including an image repository with approximately 145,000 images. Photographs were tracked, uploaded, and labeled by the research teams on a daily basis. Naturally, the labels evolved over the course of the project. With each new research team and new discovery at the site, new labels were added while others fell out of use. Researchers acknowledged the biases and subjective nature of this constantly evolving system. Modification of the labels could be challenging:

> It would be a group meeting of specialists and… they would go back and forth about what they wanted to see and at some point, either some consensus was reached, or more likely there was no consensus and someone finally said, ok I have to do something about this, and they did.
>
> The process also could privilege some research teams over others:
>
> You can see how marginalized some groups of specialists are versus others.… For some specialty you see a hundred different keyword options for a specific type of thing whereas for others you might see maybe one option for a keyword.

Thus, while there was an attempt to catalog these images alongside the excavation process using standardized keyword fields and a more or less controlled—though evolving—vocabulary, not only are the existing metadata inconsistent, approximately fifty thousand images remained without any labeling at all.

Imagine these fifty thousand images arriving in the metadata unit for cataloging. The subject analysis process would be very different from that of the domain experts. Cataloguers the authors spoke with described the approach they would take:

> We initially start with some sort of context… so in this case I would hope that we would be told that these are excavation photographs from a particular site… to narrow it down… also the timespan.

Cataloguers typically will research for details—for example, existing information about the project, its context, and history. Further questions would cover expectations and scope of the project:

> What level of detail is required to consider it adequately described.… I can see that this is an excavation site… do you want the colors

> described? …do you want the measuring thing described? …would you have an idea of what unit that is?
>
> Available time and level of effort need to be carefully balanced:
>
> Are we going image by image… do we want to organize them into groups and describe group by group[?]…
>
> There is not one practice fitting all collections or resources… so much is precluded by funding, time, and the size of the collection… when we are working with… a collection of as many as fifty thousand items… you don't have the luxury to look up every single thing… so many factors are beyond our control.
>
> That's what we are working against more and more often… the scale. We have to accept a higher rate of error….

Cataloguers may refer to existing collections elsewhere and use those catalog records as a reference but also rely on individuals with subject expertise to help them understand specific contexts. Tapping into social and institutional networks is a crucial means of extending the knowledge base for description that no one cataloguer could possibly have. But those connections are not necessarily formally identified as contributing to the work.

The discrepancy between the kind of image labelling that happens in the field and the cataloguing effort that goes on within the library is based on process, procedures, and—in the library—cataloguing rules and standards. Ultimately, however, the ability of cataloguers to meaningfully or usefully describe the images is a balance between time, their own specialized knowledge, and access to subject experts.

When Çatalhöyük researchers return to those images after they are preserved in a digital archive with the help of the library, the questions the researchers need to ask are, of course, countless and often relate to a specific research agenda. Researchers may want to use the images to determine, for example, the use of open spaces:

> I can see [in this image] that [the space] was heavily used, possibly quickly accumulated, given how much charcoal is… preserved. It wasn't immediately trampled or quickly buried so it stayed like this. It would be useful to have something like this [image] when you try to compare open spaces, but it is not often the case.

Research questions might also begin with rather specific compound queries, such as a search by keywords like *bucranium* or *burial with skeleton*; a search for images that contain a cluster or assemblages of objects, like bones combined with stone artifacts; a search for walls with painted plaster and features with particular shapes, like circular, oval, irregular holes; or a search for persons performing particular tasks like a man photographing, a woman excavating, a group touring the site, etc. The researchers' questions tend to contain within them assumptions about visual ways of perceiving characteristics of a site that are not likely to make it to the library's image description, since the act of describing, particularly by cataloguers, assumes, at best, readily recognizable qualities. The holy grail is a digital archive that not only solves the fundamental digital preservation problem of bit rot but that also makes images understandable in different contexts, that provides image description relevant to different audiences, and that can identify the image for purposes of information exchange, retrieval, and discovery.

When the Çatalhöyük image collection was accessioned, existing metadata and labels that came with the image collection served as the foundation for the application of the Metadata Object Description Schema (MODS) to prepare the collection for long-term preservation in the Stanford Digital Repository. The metadata that are of the highest interest to the researcher is all subsumed under the description field(s), making it more difficult to expose them to searches. Except for the date of the photograph and the photographer's name, all other fields are mapped onto fields labeled as notes (note 1, note 2, etc.). While MODS helps with the standardization of metadata across collections, it is less useful for a more domain-specific discovery, which becomes rather hit-or-miss. Since note fields are only available via full-text search, the researcher would have to know which terms to search for in the notes or get lucky.

Augmenting the Cataloging Process

AI—in particular, computer vision combined with machine learning—is now widely used for image classification.[9] Statistical models trained with a limited set of labeled images can, with a certain level of confidence, predict labels for unlabeled images. Given the fragmented nature of the metadata for the Çatalhöyük image collection, including a complete lack of labels for more than one-third of the collection, Engel et al. explored computer vision, using both existing image classification models as well as a simple model trained on the authors' own labels.[10]

A random sample of 766 images was offered up for labelling based on the pre-trained models provided by Clarifai and Google Vision.[11] Both of these services are trained on non-archaeological data, and consequently it is not surprising that neither is very sensitive to the identification of items within the archaeological domain. High-confidence scores for labels like *eyewear* or *subway system* demonstrate the shortcomings of the underlying models when it comes to their usefulness outside of commercial applications and contemporary objects.

The Google service proved more cautious in assigning labels. Though it rendered lower confidence and overall fewer labels per image than Clarifai, when hand-checked, Google's results were slightly more accurate than Clarifai's and reflected what appeared to be a slightly larger vocabulary related to archaeology. These differences are likely the result of different underlying training data and models. At a high level of abstraction, both services correctly assigned labels like *soil*, *archaeological site*, *geology*, and *person*.

Though these generic classes may not contribute much to discovery, we can imagine that such labels could be helpful for cataloguers.

Even when working with large, heterogeneous, unprocessed datasets, the results demonstrate that archaeological images can be distinguished from a larger pool of images. With the fifty thousand unlabeled Çatalhöyük images, a cataloguer could more easily apply collection-level subjects for the excavation and artifacts. Only with the assistance of image recognition can the cataloguer quickly identify further subsets—for example, people—making it much faster to identify individual persons or note the activities and tasks they perform.

In a second experiment, Engel et al. created a small set of training data, separating a subset of the images into four different categories.[12] An automated machine learning service was used, meaning that there was no need to do the delicate work of choosing the architecture or tuning the model; only labels and examples were required.[13] The model achieved a high percentage of correct predictions, so this approach turned out to be surprisingly useful. The limited number of classes has its shortcomings, but training the model was a very low effort, both in cost and human labor. And providing researcher-generated training data gave greater control over the model, resulting in classifications that were more domain-specific and therefore more useful.

Neither cataloguers nor archaeologists generally have the expertise or resources to train sophisticated machine models from scratch. Therefore, the availability of automated ML platforms that only require training data but take care of the work that goes into devising the best approach for training a model is very attractive.

It is obvious, however, that AI solutions cannot replace the cataloguer. Classification in the context of ML presumes that labels exist—that the classification schema has been determined already. The model is merely giving a confidence measure to whichever class is most similar. A cataloguer remains responsible for which distinctions are made. Similarly, though *object recognition* or *object detection* is the term used to describe a major breakthrough in computer vision, the model is not actually "recognizing" anything. Rather, it is a set of techniques that involves both drawing a bounding box around discrete objects in a photo and applying a class label to them. Localizing objects in this way is the result of building a collection of features based on complex pattern matching from training data. The results, while often remarkably accurate and undoubtedly useful, are not transparent and cannot be interrogated the way cataloguers challenge each other and collaborate to determine which choices to make.

Cataloguing an image involves mobilizing very different kinds of knowledge to determine what is relevant to researchers in the context of a given collection. In the process, cataloguers bring both bias and expert judgment. Different cataloguers might make different choices, depending on their respective familiarity with the context of the collection, its history, the needs of the academic community that will be studying it, the standards that will influence labelling, and so on. The cataloguers' context is multifaceted, grounded in their expertise and their accountability to their work as professionals.

Possibly one of the main reasons to introduce computation into a workflow is that a computer vision model applied to the classification task scales far better than human labor. This is particularly relevant for the above-cited problem of large collections. The nature of a computer program is to be reproducible without much effort: what could be accomplished for one image will be accomplished for an almost infinite number of other images. Cataloguers, supported by an efficient computer vision tool, would be able to significantly scale their work. It is also important to remember that models depend on the data they are trained on. Bad training data results in bad classifiers, and the cataloguers' expertise is an invaluable asset for the appropriate labeling of training data.

Conclusion: Toward Alternative Modes of Discovery

Classifications are not only subjective, they are malleable and tend to change over time. Research projects may extend over long periods during which classifications of materials are subjected to critical discussion, contested, and redefined. And yet once library collections are initially described, their description is rarely revisited or enhanced, which hinders reinterpretation. Classification in the library, as described by Broughton, is "essentially a methodology for creating an arrangement, using a particular set of criteria for grouping and ordering; provided these criteria are applied logically, and the location of any item can be reasonably predicted, we can ask no more of it than that."[14] So, if someone is looking for something about *Schubert*, that someone needs to be looking in the right place (*composers*). The need for reliable and consistent labels contributes to a hegemonic framing of information through library classification systems.[15]

An alternative approach is to move away entirely from the idea of the objective and "correct" label. Descriptions of objects, particularly images, can be multi-layered. Computer vision, while it does not provide a more objective classification, has a significant potential to enrich this multivocality by suggesting patterns that users might not have identified and providing easily calculable features like size, color, and saturation. Linked Open Data play an important role in this context. Cataloguers are beginning to make use of linked open data technology, which makes data processable by machines. It assigns unique identifiers to concepts and entities (language-based labels), which are imprecise and contestable, and become less important to interoperability and information exchange at the system level. Systems can choose which label to display and in which language, so if terminology changes over time, it is simple to update the label. Because the data can be processed by computers, previously unknown connections and relationships across institutions and domains can be revealed.[16]

Or the needs of discoverability and interoperability could be addressed separately. As Peter Leonard explained at the 2019 Fantastic Futures conference (2019),[*] "text is a sort of impoverished way of talking about the richness of visual culture." People click on "next image" or "previous image" or they go through a search bar, but the search bar involves the use of text and therefore relies on indexing, which requires prior knowledge of the

* See: https://fantasticfutures.stanford.edu.

collections, etc. But discovery might not only mean "find me X," but it could also mean "find me something that looks like *this*."[17] Images are amenable to pattern matching—measures of similarity that make clustering images, finding exact matches, and locating near-duplicates easy to do. Unsupervised ML techniques, which do not require labeled data, introduce searching based on repeating patterns, stylistics, etc.[18] Yale DH Lab's PixPlot[†] demonstrates that it is possible to navigate through a collection of images based on visual characteristics alone.

In "The visual digital turn: Using neural networks to study historical images," Wevers and Smits demonstrate multiple combinations of supervised and unsupervised approaches to reveal "visual trends" in image collections and aid in the creation of metadata.[19] Their research suggests that developing competencies in these techniques within libraries will open alternate lenses on existing visual materials. For the purpose of image discovery, similarity search can eliminate the need for a standardized text-based image description across collections and institutions. In support of cataloguing, searches across a shared platform could be conducted regularly to find duplicates in other institutions' collections and establish connections between collections that were previously unknown or to see if new entities have been identified. Combinations of these efforts can, in turn, contribute to the augmentation of the description of materials over time as users explore new ways of interpreting the collection.

In matters of artificial intelligence and starting with the very expression "artificial intelligence," it is difficult to get rid of anthropomorphism, which comes naturally to our minds. We could easily write that the computer sees something or that a computer is able to describe what it sees. But applications of AI are statistical models; they are just as subjective as the humans who build them. Determining what is relevant will continue to be the responsibility of the cataloguer and the researcher.

Endnotes

1. Joan E. Beaudoin, "Describing images: A case study of visual literacy among library and information science students," *College & Research Libraries* 77, no. 3 (2016): 376–92.
2. Mary W. Elings and Günter Waibel, "Metadata for All: Descriptive Standards and Metadata Sharing Across Cultural Heritage Collections," *VRA Bulletin* 34, no. 1 (2007): 7–14; A. Lytle, "Statewide Digitization Planners Conference Summary Report," *Microform & Imaging Review* 35, no. 1 (2006): 19–30.
3. Grace Therrell, "More product, more process: metadata in digital image collections," *Digital Library Perspectives* 35, no. 1 (2019): 2–14.
4. Margot Note, *Managing Image Collections: A Practical Guide* (Witney: Chandos Publishing, 2011), 110.
5. Lytle, "Statewide Digitization," 26.
6. Çatalhöyük Research Project, accessed April 13, 2020, http://catalhoyuk.com/.
7. Ian Hodder, "'Always Momentary, Fluid and Flexible': Towards a Reflexive Excavation Methodology," *Antiquity* 71, no. 273 (September 1997): 691–700, https://doi.org/10.1017/S0003598X00085410; see also, Ian Hodder, *The Archaeological Process: An Introduction* (Oxford, UK; Malden, MA: Blackwell, 1999); and, Ian Hodder, *Towards Reflexive Method in Archaeology: The Example at Çatalhöyük* (Cambridge: McDonald Institute for Archaeological Research, 2000).

† See: https://dhlab.yale.edu/projects/pixplot/.

8. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems* (2012), 1097–105; Peter Enser, "The evolution of visual information retrieval," *Journal of Information Science* 34, no. 4 (2008): 531–46.
9. Claudia Engel et al., "Computer Vision and Image Recognition in Archaeology" (paper presented at the Conference on Artificial Intelligence for Data Discovery and Reuse, 1–4, Pittsburgh Pennsylvania May 2019), https://doi.org/10.1145/3359115.3359117.
10. Clarifai (https://www.clarifai.com/) is an artificial intelligence company founded in 2013 specializing in computer vision. Google Cloud's Vision API (https://cloud.google.com/vision) was released in 2017.
11. Claudia Engel et al., "Computer Vision."
12. Matthias Feurer et al, "Efficient and robust automated machine learning," in *Advances in Neural Information Processing Systems* (2015), 2962–70.
13. Vanda Broughton, *Essential Classification* (London: Facet Publishing, 2015), 21.
14. Sanford Berman, *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People* (Metuchen, NJ: The Scarecrow Press, 1971); Saifya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018); Hope A. Olson, *The Power to Name: Locating the Limits of Subject Representation in Libraries* (Dordrecht: Springer Science & Business Media, 2013).
15. Christian Bizer, Tom Heath, and Tim Berners-Lee, "Linked Data—The Story So Far," *International Journal on Semantic Web and Information Systems (IJSWIS)* 5, no. 3 (2009): 1–22; Philip Evan Schreur, "The academy unbound," *Library Resources & Technical Services* 56, no. 4 (2012): 227–37.
16. "Fantastic Futures," Stanford Libraries, https://fantasticfutures.stanford.edu.
17. Peter Leonard, "Collection-scale visualization with PixPlot," filmed December 4, 2019, in Stanford, CA, United States of America, Fantastic Futures 2019, https://youtu.be/3lREJYBLRnY.
18. Junyuan Xie, Ross Girshick, and Ali Farhadi, "Unsupervised deep embedding for clustering analysis," in *International Conference on Machine Learning* (2016), 478–87.
19. Melvin Wevers and Thomas Smits, "The Visual Digital Turn: Using Neural Networks to Study Historical Images," *Digital Scholarship in the Humanities* 35, no. 1 (2020): 194–207, https://doi.org/10.1093/llc/fqy085. An important precursor to this work is Isabella di Leonard, Benoît Laurent Auguste Seguin, and Frédéric Kaplan, "Visual Patterns Discovery in Large Databases of Paintings," *Infoscience* (2016), https://infoscience.epfl.ch/record/220638.

Bibliography

Beaudoin, Joan E. "Describing images: A case study of visual literacy among library and information science students." *College & Research Libraries* 77, no. 3 (2016): 376–92. https://doi.org/10.5860/crl.77.3.376.

Berman, Sanford. *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People*. Metuchen, NJ: The Scarecrow Press, 1971.

Bizer, Christian, Tom Heath, and Tim Berners-Lee. "Linked Data—The Story So Far." *International Journal on Semantic Web and Information Systems (IJSWIS)* 5, no. 3 (2009): 1–22.

Broughton, Vanda. *Essential Classification*. London: Facet Publishing, 2015.

Çatalhöyük Research Project. Accessed April 13, 2020. http://catalhoyuk.com/.

Clarifai. 2020. https://www.clarifai.com/.

Elings, Mary W., and Günter Waibel. "Metadata for All: Descriptive Standards and Metadata Sharing Across Cultural Heritage Collections." *VRA Bulletin* 34, no. 1 (2007): 7–14.

Engel, Claudia, Peter Mangiafico, Justine Issavi, and Dominik Lukas. "Computer Vision and Image Recognition in Archaeology." Paper presented at the Conference on Artificial Intelligence for Data Discovery and Reuse, 1–4. Pittsburgh Pennsylvania May 2019. https://doi.org/10.1145/3359115.3359117.

Enser, Peter. "The evolution of visual information retrieval." *Journal of Information Science* 34, no. 4 (2008): 531–46.

Feurer, Matthias, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. "Efficient and robust automated machine learning." In *Advances in Neural Information Processing Systems*, 2962–70, 2015.

Hodder, Ian. "'Always Momentary, Fluid and Flexible': Towards a Reflexive Excavation Methodology." *Antiquity* 71, no. 273 (September 1997): 691–700. https://doi.org/10.1017/S0003598X00085410.
Hodder, Ian. *The Archaeological Process: An Introduction*. Oxford, UK; Malden, MA: Blackwell, 1999.
———. *Towards Reflexive Method in Archaeology: The Example at* Çatalhöyük. Cambridge: McDonald Institute for Archaeological Research, 2000.
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in Neural Information Processing Systems*, 1097–105, 2012.
Lenardo, Isabella di, Benoît Laurent Auguste Seguin, and Frédéric Kaplan. "Visual Patterns Discovery in Large Databases of Paintings." *Infoscience* (2016). https://infoscience.epfl.ch/record/220638.
Leonard, Peter. "Collection-scale visualization with PixPlot." Filmed December 4, 2019, in Stanford, CA, United States of America. Fantastic Futures 2019. https://youtu.be/3lREJYBLRnY.
Lytle, A. "Statewide digitization planners conference summary report." *Microform & Imaging Review* 35, no. 1 (2006): 19–30.
Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.
Note, Margot. *Managing Image Collections: A Practical Guide*. Witney: Chandos Publishing, 2011.
Olson, Hope A. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht: Springer Science & Business Media, 2013.
RDA Toolkit. American Library Association, Canadian Federation of Library Associations, and CILIP: Chartered Institute of Library and Information Professionals, 2020. http://access.rdatoolkit.org/.
Schreur, Philip Evan. "The academy unbound." *Library Resources & Technical Services* 56, no. 4 (2012): 227–37.
Stanford Libraries. "Fantastic Futures." Accessed May 25, 2020. 2019. https://fantasticfutures.stanford.edu.
Therrell, Grace. "More product, more process: metadata in digital image collections." *Digital Library Perspectives* 35, no. 1 (2019): 2–14.
Vision AI. Google Cloud. 2020. https://cloud.google.com/vision/.
Wevers, Melvin, and Thomas Smits. "The Visual Digital Turn: Using Neural Networks to Study Historical Images." *Digital Scholarship in the Humanities* 35, no. 1 (2020): 194–207. https://doi.org/10.1093/llc/fqy085.
Xie, Junyuan, Ross Girshick, and Ali Farhadi. "Unsupervised deep embedding for clustering analysis." In *International Conference on Machine Learning*, 478–87. 2016.
Yale University Digital Humanities Lab. "PixPlot." Accessed July 26, 2020. https://dhlab.yale.edu/projects/pixplot/.

Chapter 8

AI-Informed Approaches to Metadata Tagging for Improved Resource Discovery

Charlie Harper, Anne Kumer, Shelby Stuart, and Evan Meszaros

Introduction

Academic and cultural institutions are grappling with problems of how to organize, label, and search disparate bodies of texts. As aggregators, preservers, and disseminators of substantial repositories of digital texts, research libraries are naturally situated at the heart of these problems. This chapter explores how unsupervised machine learning may be used to capture and simplify the complexity and nuances of text. Traditional approaches to improving discoverability and accessibility of text through metadata and controlled vocabularies have time-tested strengths. As the volume of digital data explodes, the obstacles and limitations of traditional approaches become more pronounced, and machine learning "show(s) the potential to create efficiencies that smooth the path to access, enhancing description and expanding forms of discovery along the way."[1] In light of the need for new approaches to metadata generation to facilitate discovery, the authors look at Doc2Vec and topic modelling with Latent Dirichlet Allocation (LDA) to explore their utility as assistive tools for authors, librarians, and readers. The authors apply the two approaches to a corpus of electronic theses and dissertations (ETDs) completed at Ohio universities and colleges.*

Current Issues in Metadata and Discovery

Searchability is one of the greatest advantages that online documents have over their print counterparts, and surveys show that users view this as a vital feature when asked about using e-resources over print.[2] Metadata quality influences the searchability and the discoverability of e-resources. Research databases and discovery layers rely on proprietary algorithms to generate and order results in response to the user's query. Relevance ranking algorithms may compare the query to metadata fields such as subject headings, publication titles, abstracts, and (sometimes) full text in order to determine the results. Therefore, search engines will return resources with greater effectiveness and precision when they have complete metadata and a useful set of subject headings. High-quality metadata is also a key component in ensuring that the most relevant documents appear at the top of the result set, where the user is more likely to find them.[3]

Studies by Tina Gross and her colleagues have examined the efficacy of controlled vocabularies for resource discovery. They established that, whether or not a user sees them, the existence of controlled vocabulary metadata, which depends on carefully assigned subject headings, generally contributes to up to one-third more positive search results than if that metadata was not there.[4] The research landscape, however, has changed dramatically due to Google's powerful influence, and keyword searching has exploded in popularity. The millions of documents that are commonly returned by keyword searches may overwhelm the user, while subject searches are able to return smaller sets of documents that are often more tailored to a user's query. Concurrently, several LIS scholars find that unregulated author-generated keywords enhance searches if they are employed in addition to subject headings from widely used controlled vocabularies assigned by librarians.[5]

The most widely used library-controlled vocabulary, the Library of Congress Subject Headings (LCSH), is maintained on the principle of literary warrant.[6] This has historically meant that only topics published in books warrant inclusion in the vocabulary's authorized headings lists. Vocabularies like the LCSH are slow to add new, potentially dubious terminology, essentially "controlling" its terms by applying parameters for use. This principle neglects formats, such as articles and dissertations, where scholarship is typically first published.[7] A contrasting principle is user warrant, which is based on user preference, need, and search patterns. Leaving out the specialized knowledge of a document's author potentially lessens discoverability because the LCSH is slow to include new specialized subject terminology. ANSI/NISO standards present literary and user warrant as complementary and equally important for search and discovery.[8] Author-generated keywords may yield many irrelevant search results, which the restriction of a controlled

* This study's data sets, python notebooks, and trained models are provided on OSF (https://osf.io/r6yhp/) and are licensed under Creative Commons Attribution-ShareAlike 4.0.

vocabulary mitigates. Conversely, a controlled vocabulary imposes conservatism in the face of shifting cultural standards, which is balanced by author-generated keywords.

ETDs and Subject Metadata

For many universities and colleges, the transition from print to electronic theses and dissertations began in the mid-2000s. With this format change, librarians were able to harvest author-supplied keywords from the electronic submission forms that accompanied ETDs and include those in the dissertation's catalog record alongside cataloger-supplied descriptive subject headings to enhance search and discovery. When selecting keywords, authors tend to choose those that represent their experiences and expectations rather than terms that derive from "any kind of controlled indexing language or authority-controlled procedure."[9] Personal experiences and social motivators, such as opinion, expression, performance, and activism, can impact keyword choice and result in both overly specific and overly broad keywords.[10] As Yelton notes for MIT's ETD repository, "Most of [the author-assigned keywords] are so granular that they apply to only one thesis and therefore don't collocate anything."[11] An ETD cataloged with only highly specialized or overly broad keywords does little to enhance search and discovery.

At the same time, ETDs are particularly important when researching topics that are new and emerging. McCutcheon notes that while print theses and dissertations tend to receive little attention, "it's not uncommon for ETDs to be downloaded hundreds or thousands of times, from all over the world."[12] As gray literature, however, ETDs do not benefit from the kinds of support that are offered by commercial publishers. They lack, for instance, standard distribution channels and presence on major publishers' web platforms. In addition, ETDs are not necessarily indexed by major abstracting and indexing services, which can make them difficult to discover. ETDs are accordingly a prime dataset for projects that aim to improve metadata and increase discoverability. In order to address this problem, the authors elected to work with ETDs published at Ohio colleges and universities. These ETDs are hosted by OhioLINK,[†] a consortium of over one hundred academic institution members across the state of Ohio, and they have consumable metadata available through the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).[‡] ETDs published on OhioLINK are globally accessible, free of charge, and frequently include born-digital PDFs.

The authors wrote a series of Python notebooks to generate a dataset of OhioLINK ETDs. First, the authors used Python's Sickle library to pull Dublin Core metadata for ETDs that were published and uploaded in 2019. From the Dublin Core XML results, the authors created one CSV of the title, abstract, publication date, source university/college, URI, and rights restrictions, as well as a second CSV of the keywords assigned to each ETD. The final dataset consisted of metadata for 3,316 ETDs from thirty-six Ohio universities and colleges and 13,141 non-unique keywords.

---

† See: https://www.ohiolink.edu/.

‡ See https://www.openarchives.org/pmh/.

Representation of the thirty-six Ohio universities and colleges was highly uneven within the dataset. For example, Ohio State University produced 843 ETDs, while smaller institutions produced only one. The different academic focuses of each institution likely means that the subject areas of the dataset are skewed. Keywords that occur over one hundred times give a sense of how the subjects trend (table 8.1). Since 85.88 percent (11,285) of keywords occur only once, however, this list should be read cautiously. Likewise, the length of the abstracts is highly varied, which may further bias the dataset toward particular subject areas.

Table 8.1
Keywords that occur more than 100 times in the dataset of 3,316 ETDs. These keywords hint at how the content of the dataset may be skewed toward certain subjects.

Keyword	Occurrences
psychology	220
biology	175
education	169
mechanical engineering	154
chemistry	134
electrical engineering	133
computer science	128
communication	107
literature	106

Tagging ETDs with Doc2Vec and DBPedia

Doc2Vec is an approach that learns to map units of text, such as sentences, paragraphs, or full documents, into a numerical vector space.[13] It is an extension of an earlier, and still frequently used, incarnation known as Word2Vec, which worked with single words.[14] Both Word2Vec and Doc2Vec are built on a neural network architecture that trains on a corpus of text and learns how to represent text as coordinates in a high-dimensional space.[15] The value of these learned coordinates is that the topology of the vector space in which the text is embedded holds information on the content or meaning of the text. For example, embedded texts that are located more closely should also show a closer semantic relationship. Mathematical connections between points can also reveal deeper linguistic relationships. With single words, one can discover antonyms, synonyms, declensions, or conjugations based on spatial relationships (figure 8.1). Doc2Vec extends Word2Vec's capabilities to texts of any length.

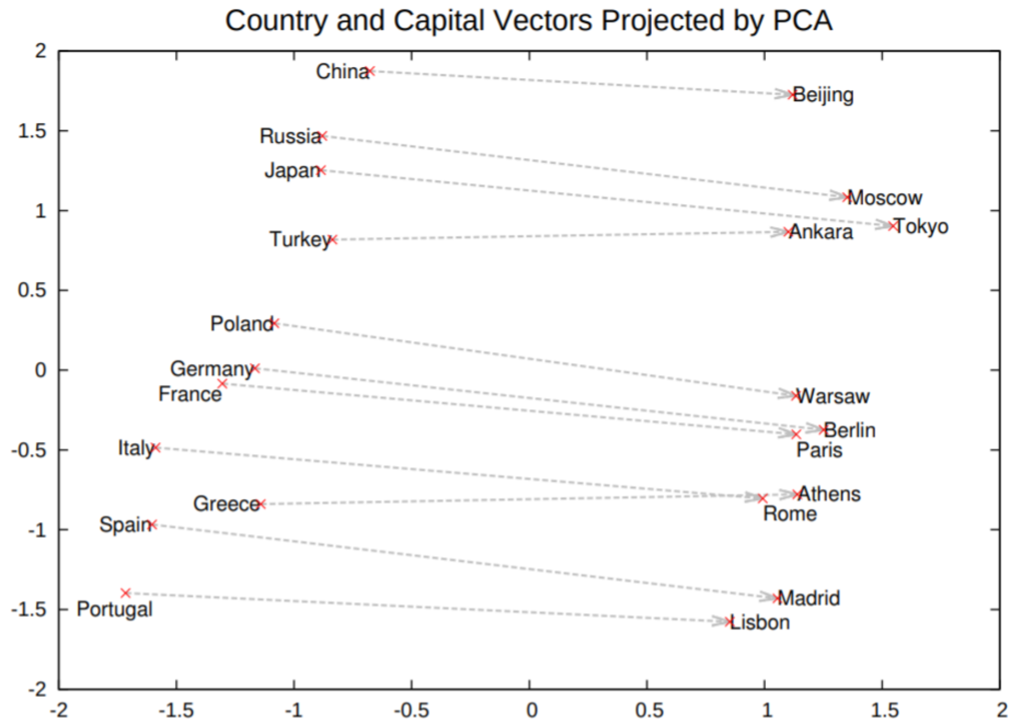


Figure 8.1
A classic example of how Word2Vec can capture meaning is the relationship between capital cities and countries learned from a corpus of text. The spatial relationship between the learned word embeddings for country and capital reflects the semantic relationships between the words in text.[16]

Doc2Vec has shown particular application in document retrieval systems, where a user can search for documents whose content or subject is related to an input document. In the library world, Yelton used Doc2Vec in her app, HAMLET, to calculate the similarity between graduate theses at MIT.[17] As Yelton notes, however, Doc2Vec cannot assign meaningful labels to related documents in the traditional sense of metadata keywords or subject headings, nor can it draw boundaries to create discrete categories of documents.[18] This is part of a larger issue with unsupervised machine learning, which reveals similarities in data but still requires humans to assign meaningful labels or keywords. In order to overcome this limitation and to automatically generate content-specific keywords, the authors trained Doc2Vec on a corpus of text generated from DBPedia, a large linked and already-labeled dataset.[19] The authors then tagged a sample of OhioLINK ETDs with the trained model to assess its effectiveness.

DBPedia and Model Training

DBPedia* is a knowledge base that classifies content using descriptive terms as well as the contextual relationships of its content. As a source for descriptive keywords, DBPedia has multiple strengths: it is crowdsourced and likely to remain more current than controlled vocabularies; its entries are internally linked to enhance semantic queries; it provides a URI, keyword, and abstract for each idea; its keywords are frequently multilingual; and a single abstract and URI can map to multiple keywords that capture the same idea.

The authors used Python's SPARQLWrapper library† to gather three hierarchical levels of data from DBPedia's SPARQL endpoint, which the authors termed page-level, subject-level, and concept-level. Page-level data is the finest grained and maps to a single entry with an abstract. Subject-level data is marked by the RDF verb "dct:subject-of" and aggregates related page-level data. Concept-level data is marked by the RDF verb "skos:broader-of" and aggregates subject-level data. Neither subjects nor concepts possess abstracts. The three should respectively represent a continuum from more specific to more general ideas (figure 8.2).

The DBPedia dataset consisted of 4,935,271 pages.‡ Abstracts ranged from 1 to 168,193 words with an average of 525 words. Initial experiments with the entire body of abstracts showed poor results, which the authors attributed to the prevalence of shorter abstracts that did not convey enough meaning. Therefore, the authors removed all but the 75th through 99.9th percentile of abstracts based on word count. The authors felt the resulting range of 648 to 5,127 words was more reasonable. This subset of 1,230,980 abstracts constituted the training set for the Doc2Vec model.

The authors used Python's Gensim library to build the Doc2Vec model.[20] Because model accuracy can be difficult to measure in unsupervised learning, the past work on Doc2Vec with Wikipedia, the computational time for training, and the authors' interpretation of experimental results guided hyperparameter choices.[21] Ultimately, the authors chose to use a continuous bag of words with a vector space of 500 dimensions. DBPedia abstracts were preprocessed by removing non-alphanumeric characters, stopwording, and lemmatizing. Training took approximately 2.5 hours on an Amazon Web Services (AWS) r5.4xlarge instance. After training, a k-d tree was built from the embedded page vectors stored in the Doc2Vec model in order to speed the search for the closest (measured by Euclidean§ distance) points in 500 dimensions.[22]

To test the efficacy of this approach, the authors tagged a selection of ETD abstracts with the page-level keywords that were closed in vector space. Tagging was accomplished by first embedding an ETD's preprocessed abstract in 500-dimensional space with the trained Doc2Vec model and then searching the k-d tree for the n-nearest points, each of

* See https://wiki.dbpedia.org/.

† See https://rdflib.dev/sparqlwrapper/.

‡ This study employed the DBPedia version 2016–10 release for page-level metadata and abstracts (https://wiki.dbpedia.org/develop/datasets/dbpedia-version-2016-10).

§ Euclidean distance extends the measure of distance as expressed in the Pythagorean Theorem to n-dimensions.

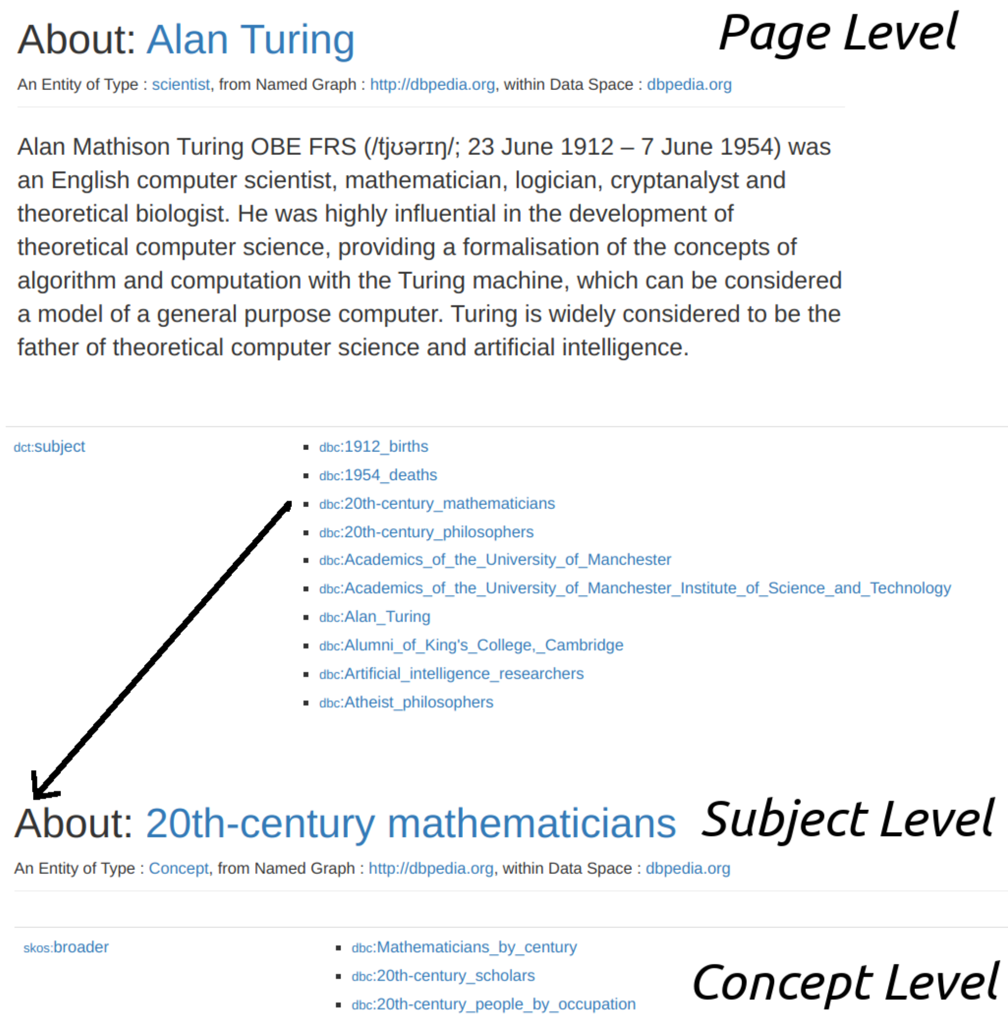


Figure 8.2
An example of a partial page with abstract (http://dbpedia.org/page/Alan_Turing) and a linked subject (http://dbpedia.org/page/Category:20th-century_mathematicians) with reference to its higher concepts.

which corresponds to one DBPedia page. The results were extremely poor and typically nonsensical. For example, one thesis on college students' perceptions of conservation efforts was tagged "Keg_stand"! The authors concluded that the information contained at the page level was overly specific and that the vector space was likely too densely packed with points. To overcome this, the authors utilized the linked nature of DBPedia to move up to the subject ("dct:subject") and concept ("skos:broader") levels for tagging.

The subset of 1,230,980 abstracts linked to 728,752 subjects and 421,051 concepts. Subjects mapped to a range of 1 (e.g., "Crocodile_Dundee_Films") to 177,622 ("Living_People") page-level abstracts. Concepts mapped to a range of 1 (e.g., "1130s_in_Europe") to 5,063 ("Songs_by_songwriter") subjects. Because of the interlinked nature of DBPedia, there is overlap between subject and concept keywords. To build a k-d tree for the subject level, the vectors of each subject's pages were averaged together. For the concept level, the vectors of each concept's subjects were averaged together (figure 8.3). The trained Doc2Vec model was unaltered.

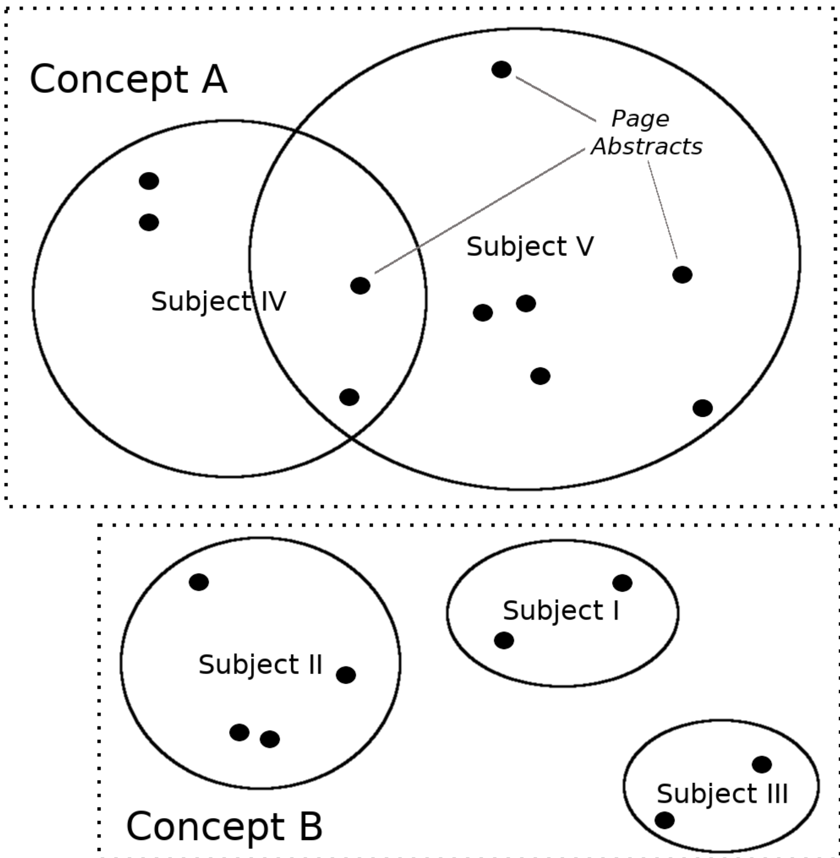


Figure 8.3
Illustration of moving from page- to subject- to concept-level in the vector space using relationships stored in DBPedia. For example, Subject IV contains four pages with abstracts, represented by black dots. These four points, which in reality are 500-dimensional, are averaged together to create Subject IV, a new, 500-dimensional point. To create Concept A, Subject IV and Subject V are averaged together. As one moves from page to concept, the vector space becomes sparser and content should become more general. Note that pages can belong to multiple subjects, and subjects can belong to multiple concepts.

In order to test this approach for subject and concept tagging, the authors sampled 250 ETDs published in 2019. The sample was stratified by university/college in order to reflect the distribution of institutions in OhioLINK. The Doc2Vec model was used to embed each ETD's preprocessed abstract into vector space and then the subject and concept k-d trees were searched to find the five nearest subjects and concepts as measured by Euclidean distance (table 8.2).

Table 8.2
An example of subject and concept DBPedia tags assigned to an ETD entitled, "Development of a Conformal Additive Manufacturing Process and its Application."[23]

	1	2	3	4	5
Subject	Nanotechnology	Materials_ science	Lithography_ (microfabrication)	Microtechnology	Semiconductor_ device_ fabrication
Concept	Microtechnology	Computer-aided_ engineering	Materials_ science	Forming_ processes	Instrumental_ analysis

Results

The individual authors each rated 125 ETD's subject and concept tags to ensure that tags were always rated by two separate individuals. For simplicity, each rater marked the relevance of the tag -1 (not relevant), 0 (somewhat relevant), or 1 (relevant). The ratings were then averaged across raters. Averaged ratings for subjects were more relevant, on average, than for concepts. In both cases, moving from the first subject or concept (closest in space) to the fifth subject or concept (farther in space) showed a downward trend in ratings (table 8.3).

The mean subject rating was 0.32548 ±0.057. The mean concept rating was 0.23496 ±0.057. Subjects and concepts were, therefore, both ranked as being "somewhat relevant" on the whole to the ETDs. This result is far from perfect, but it is very promising. While page-level tagging produced no meaningful results, at the subject and concept level, this approach is capturing meaning and assigning viable keywords based only on an abstract.

Table 8.3
The mean and 95% confidence interval for subject and concept ratings based on a sample of 250 tagged ETDs.

	Mean	95% Lower	95% Upper
subject1	0.4630	0.379	0.547
subject2	0.3817	0.294	0.469
subject3	0.3471	0.261	0.433
subject4	0.2396	0.155	0.324
subject5	0.1942	0.102	0.286
concept1	0.3104	0.230	0.390
concept2	0.3389	0.257	0.420

Table 8.3
The mean and 95% confidence interval for subject and concept ratings based on a sample of 250 tagged ETDs.

	Mean	95% Lower	95% Upper
concept3	0.1925	0.102	0.282
concept4	0.1958	0.105	0.287
concept5	0.1208	0.035	0.207
subject_avg	0.32548	0.2684	0.3825
concept_avg	0.23496	0.1785	0.2915

Finding Relevant Information with Topic Modeling

Topic modeling is a generative statistical approach that clusters related content. This term is commonly a stand-in for the more specific topic modeling algorithm, Latent Dirichlet Allocation, or LDA.[24] The approach is often employed in fields that engage with large corpora of textual data. In the academic library, researchers have already used topic modeling to cluster ETDs and government documents for enhanced discovery to generate alt-metrics by mining book reviews and to recommend tags for enhancing metadata records.[25]

Despite its value in certain applications, there are notable shortcomings with topic modeling. Foremost, "topic" is a misnomer. As a statistical method, LDA produces a statistical distribution of words that constitute a "topic" and a statistical distribution of "topics" across documents. Often, scholars will choose the top *n* words to represent a topic, but LDA does not produce a label for a topic, nor does it guarantee the top n words are meaningful to a human reader. Second, LDA requires a preset number of topics. There are methods to best determine this, but if a trained model is continuously applied to a naturally growing corpus, such as is the case with ETDs, the number of topics is unable to organically grow with the changing content.

For these reasons, the authors believe that topic modeling retains immense use for clustering fixed corpora of text but that it is less useful for a living corpus. While an approach like the combined Doc2Vec and DBPedia above is best situated to generate metadata to improve the discovery of resources within a large, living corpus of ETDs, topic modeling is better suited to enhance discovery of specific information within an ETD, which is, in effect, a fixed corpus.

ETD Full Text and Model Training

To exemplify the authors' proposal that LDA is most useful for internal information discovery, topic models were trained on the full text of ten ETDs from the previous sample of 250. The full text was extracted from each PDF in Python. Because of difficulties in working with non-standardized ETDs, the authors chose the page as the basic unit of analysis when training the topic models. No other preprocessing was done.

An LDA model was trained on each ETD's set of pages using the Gensim library. The number of topics was set at ten, which seemed reasonable to capture enough nuance in ETDs of variable length. The model used only words that appeared on at least five pages but fewer than 25 percent of pages. After training, a CSV of topic distributions for each page was generated and the top five words for each topic were stored. The LDA model was then discarded.

Results

Assessing the results of topic models is difficult and requires specialized knowledge and deeper engagement with each ETD's content. Visualizing the results, however, does show the strong potential of this approach for assisting readers in finding information within an ETD. Figure 8.4 shows the visualization of topic distributions by page for an MA thesis entitled, "Enduring Failure: A Borderlands History of the Iraq War and its Aftermath."[26] Without hyperparameter tuning, the LDA model has produced generally good topics. The fifth topic, "general, saddam, intelligence, regime, bush," is an example of this. The topic is absent from the first portion of the text and clusters around pages in the fifties and sixties. If a reader were interested in the rhetoric, personalities, and intelligence that led up to the Iraq War, this would indicate that the reader should glance at these pages first.

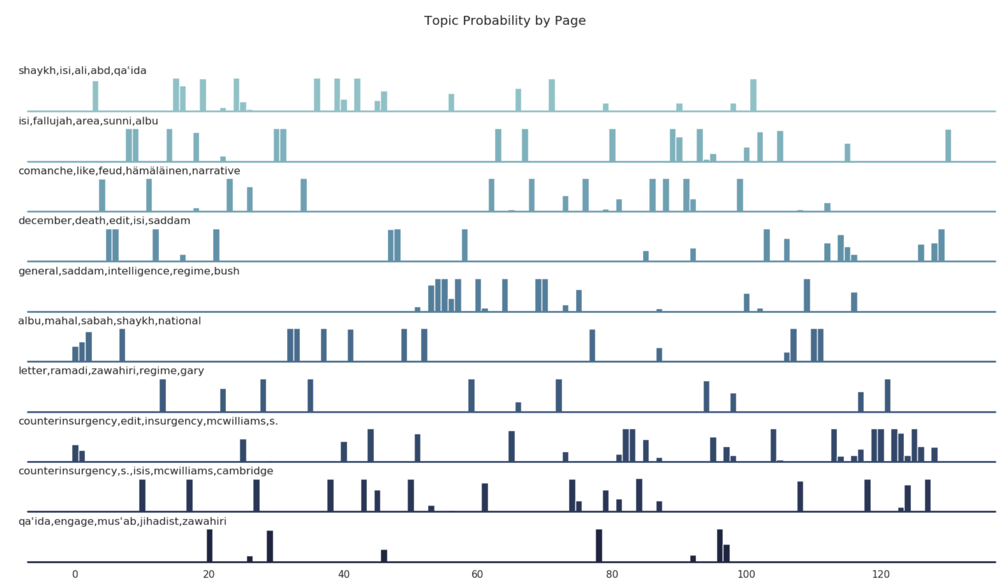


Figure 8.4

The distribution of topics across the pages of an ETD on the Iraq War.[27] Ten topics are presented, from top to bottom, with the top five words for each topic. The mixture of each topic by page is shown from left to right.

Conclusion

A Doc2Vec model trained on DBPedia's linked content and topic models trained on individual ETDs show promise as tools to enhance metadata and discovery. Both approaches outlined above warrant deeper study and the authors are pursuing ways to improve and better assess their efficacy. Regardless, these approaches seem well-poised to inform human metadata creation and discovery efforts but not to replace them. Although the Doc2Vec subject and concept tags were generally relevant to the ETDs' abstracts, there is substantial room for improvement and model tuning. In addition, finding ways to better tune topic models to individual ETDs would produce stronger results. In the course of this work, the authors made numerous observations that are guiding their ongoing work. Many observations additionally reflect deeper issues with the rising tide of machine learning in the library. Although only a handful of these can be enumerated here, the authors find it beneficial to conclude with the following:

1. It is difficult to judge model effectiveness. Rating machine-generated tags and topics require a baseline level of subject expertise and familiarity with terminology, which is especially important when documents in the sample set have been written by and for graduate-level researchers. Of the authors, those who had educational backgrounds in the social sciences and humanities struggled to assess the relevance of some tags assigned to, for instance, physics and engineering ETDs. It is, therefore, advisable to engage with subject-matter experts when assessing the effectiveness of machine learning approaches to tagging and discovery.
2. Linked data augment discovery. One oft-noted benefit of employing controlled subject headings is that they integrate the ETDs with materials that share the same subject but are published in different formats. This increases the visibility of the ETDs, which otherwise may only be retrievable by searching within a particular repository or library collection and exposes them to a much broader range of researchers.[28] Utilizing keywords drawn from DBPedia's linked data set may offer an additional way to interlink ETDs with other academic resources. Moreover, following links between keywords may facilitate the sort of serendipitous discovery that can occur when browsing print items on a library shelf.
3. All subjects are not created equal. Abstracts for humanities ETDs, such as those describing poetry collections, creative writing, theater productions, and others, were less likely to be assigned relevant tags. This could be related to the tendency of those abstracts to have smaller word counts than their STEM counterparts. Moreover, the authors observed a lack of accuracy and specificity in tagging ETDs that examine certain understudied communities and locations. Among the sampled ETDs, this issue seemed particularly common among those that focused on Latin America. For example, an ETD studying public performances in Colombia was tagged "Argentine Art," and one describing ecological research in the Peruvian Andes was tagged "Forestry in Brazil." As mentioned previously, the ETDs are likely biased toward certain subject areas as are the DBPedia abstracts.

These biases in datasets become reified in machine learning models and can contribute to results that show an even stronger bias.

4. Humans and machines need balance. Authors choose keywords from a place of ownership and perceived use of their scholarship, librarians apply subject headings in compliance with best practices and parameters for metadata quality control, and machine learning models select terms or topics according to patterns learned from human-supplied data. No one method is ideal, and a balance between the strengths and weaknesses of each is needed; the human capability to shift perspective and interpret words or phrases in different contexts is not directly replicated by machine learning methods, while a machine learning model's ability to rapidly process huge corpora cannot be directly replicated by a human. Mediating the differing roles and biases of author, librarian, and machine requires ongoing research and human devotion to consistency. Cataloging best practices remains essential for quality control when applying machine learning techniques to resource description.

Endnotes

1. Thomas Padilla, *Responsible Operations: Data Science, Machine Learning, and AI in Libraries* (Dublin, OH: OCLC Research, 2019), 12, https://doi.org/10.25333/xk7z-9g97.
2. Gabrielle Wiersma and Esta Tovstiadi, "Inconsistencies between Academic E-Book Platforms: A Comparison of Metadata and Search Results," *portal: Libraries and the Academy* 17, no. 3 (2017): 628.
3. Robert Losee, "The Effect of Assigning a Metadata or Indexing Term on Document Ordering," *Journal of the American Society for Information Science and Technology* 64, no. 11 (2013).
4. Tina Gross, Arlene G. Taylor, and Daniel N. Joudrey, "Still a Lot to Lose: The Role of Controlled Vocabulary in Keyword Searching," *Cataloging & Classification Quarterly* 53, no. 1 (2015): 1–39, https://doi.org/10.1080/01639374.2014.917447.
5. Tom Steele and Nicole Sump-crethar, "Metadata for Electronic Theses and Dissertations: A Survey of Institutional Repositories," *Journal of Library Metadata* 16, no. 1 (2016): 53–68, https://doi.org/10.1080/19386389.2016.1161462; Gross, Taylor, and Joudrey, "Still a Lot to Lose"; Sevim McCutcheon, "Basic, Fuller, Fullest: Treatment Options for Electronic Theses and Dissertations," *Library Collections, Acquisitions, & Technical Services* 35, no. 2–3 (2011): 64–68, https://doi.org/10.1016/j.lcats.2011.03.019.
6. Cynthia R. Strader, "Author-Assigned Keywords Versus Library of Congress Subject Headings," *Library Resources and Technical Services* 53, no. 4 (2011): 243–51, https://doi.org/10.5860/lrts.53n4.243.
7. Strader, "Author-Assigned Keywords."
8. Ibid.
9. Michalis Gerolimos, "Tagging for Libraries: A Review of the Effectiveness of Tagging Systems for Library Catalogs," *Journal of Library Metadata* 13, no. 1 (2013): 41, https://doi.org/10.1080/19386389.2013.778730.
10. Alla Zollers, "Emerging Motivations for Tagging: Expression, Performance, and Activism," in *WWW 2007: Proceedings of the 16th International Conference on World Wide Web, Banff, (Alberta, Canada), May 8-12, 2007*, ed. Carrie Williamson and Mary E. Zurko (New York: ACM Press, 2007). http://wwwconference.org/www2007/workshops/paper_55.pdf.
11. Andromeda Yelton, "HAMLET: Neural-Net-Powered Prototypes for Library Discovery," in *Artificial Intelligence and Machine Learning in Libraries*, ed. Jason Griffey (Chicago: ALA TechSource, 2019), 12.
12. McCutcheon, "Basic, Fuller, Fullest," 64.

13. Quoc V. Le and Tomas Mikolov, "Distributed Representations of Sentences and Documents," in *Proceedings of the 31st International Conference on Machine Learning, PMLR* 32 (2014): 1188–96, https://arxiv.org/abs/1405.4053.
14. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean, "Efficient Estimation of Word Representations in Vector Space," Cornell University (2013), 1–12, https://arxiv.org/abs/1301.3781.
15. Hobson Lane, Hannes M. Hapke, and Cole Howard, *Natural Language Processing in Action: Understanding, Analyzing, and Generating Text with Python* (Shelter Island, NY: Manning Publications, 2019), 181–217.
16. Mikolov et al., "Distributed Representations of Words," fig. 2.
17. Yelton, "HAMLET."
18. Ibid., 11–12.
19. M. Allahyari and K. Kochut, "Semantic Tagging Using Topic Models Exploiting Wikipedia Category Network," in *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)* (New York: IEEE Xplore, 2016), 63–70, https://doi.org/10.1109/ICSC.2016.34.
20. Radim Řehůřek and Petr Sojka, "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (Valletta, Malta: ELRA, 2010), 45–50.
21. Jey H. Lau and Timothy Baldwin, "An Empirical Evaluation of Doc2vec with Practical Insights into Document Embedding Generation," *Proceedings of the 1st Workshop on Representation Learning for NLP* (2016), 78–86.
22. G. Bonaccorso, *Hands-On Unsupervised Learning with Python: Implement Machine Learning and Deep Learning Models Using Scikit-Learn, TensorFlow, and More* (Birmingham: Packt Publishing, 2019), 66–71.
23. Faez Alkadi, "Development of a Conformal Additive Manufacturing Process and Its Application" (MA Thesis, University of Akron, 2019).
24. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research* 3 (2003): 993–1022.
25. Yelton, "HAMLET"; Jonathan Cain, "Using Topic Modeling to Enhance Access to Library Digital Collections," *Journal of Web Librarianship* 10, no. 3 (2016); Zhou Qingqing and Zhang Chengzhi, "Measuring Book Impact via Content-Level Academic Review Mining," *The Electronic Library* 38, no. 1 (2020): 138–54, https://doi.org/10.1108/EL-08-2019-0184; Suppawong Tuarob, Line C. Pouchard, and C. Lee Giles, "Automatic Tag Recommendation for Metadata Annotation Using Probabilistic Topic Modeling," in *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '13* (New York: Association for Computing Machinery, 2013), 239–48, https://doi.org/10.1145/2467696.2467706.
26. Terry T. Tait, "Enduring Failures: A Borderlands History of the Iraq War and Its Aftermath" (MA Thesis, Miami University, 2019).
27. Tait, "Enduring Failures."
28. McCutcheon, "Basic, Fuller, Fullest."

Bibliography

Alkadi, Faez. "Development of a Conformal Additive Manufacturing Process and Its Application." MA Thesis, University of Akron, 2019.

Allahyari, M., and K. Kochut. "Semantic Tagging Using Topic Models Exploiting Wikipedia Category Network." In *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, 63–70. New York: IEEE Xplore, 2016. https://doi.org/10.1109/ICSC.2016.34.

Blei, D. M., A. Y. Ng, and M. I. Jordan. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (2003): 993–1022.

Bonaccorso, G. *Hands-On Unsupervised Learning with Python: Implement Machine Learning and Deep Learning Models Using Scikit-Learn, TensorFlow, and More*. Birmingham: Packt Publishing, 2019.

Cain, Jonathan. "Using Topic Modeling to Enhance Access to Library Digital Collections." *Journal of Web Librarianship* 10, no. 3 (2016).

Gerolimos, Michalis. “Tagging for Libraries: A Review of the Effectiveness of Tagging Systems for Library Catalogs.” *Journal of Library Metadata* 13, no. 1 (2013): 36–58. https://doi.org/10.1080/19386389.2013.778730.

Gross, Tina, Arlene G. Taylor, and Daniel N. Joudrey. “Still a Lot to Lose: The Role of Controlled Vocabulary in Keyword Searching.” *Cataloging & Classification Quarterly* 53, no. 1 (2015): 1–39. https://doi.org/10.1080/01639374.2014.917447.

Lane, Hobson, Hannes M. Hapke, and Cole Howard. *Natural Language Processing in Action: Understanding, Analyzing, and Generating Text with Python*. Shelter Island, NY: Manning Publications, 2019.

Lau, Jey H., and Timothy Baldwin. “An Empirical Evaluation of Doc2vec with Practical Insights into Document Embedding Generation.” *Proceedings of the 1st Workshop on Representation Learning for NLP* (2016), 78–86.

Le, Quoc V., and Tomas Mikolov. “Distributed Representations of Sentences and Documents.” In *Proceedings of the 31st International Conference on Machine Learning, PMLR 32*, 2 (2014), 1188–96. https://arxiv.org/abs/1405.4053.

Losee, Robert. “The Effect of Assigning a Metadata or Indexing Term on Document Ordering.” *Journal of the American Society for Information Science and Technology* 64, no. 11 (2013).

McCutcheon, Sevim. “Basic, Fuller, Fullest: Treatment Options for Electronic Theses and Dissertations.” *Library Collections, Acquisitions, & Technical Services* 35, no. 2–3 (2011): 64–68. https://doi.org/10.1016/j.lcats.2011.03.019.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. “Efficient Estimation of Word Representations in Vector Space,” Cornell University (2013), 1–12. https://arxiv.org/abs/1301.3781.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, G. S. Corrado, and Jeffrey Dean. “Distributed Representations of Words and Phrases and Their Compositionality.” *Advances in Neural Information Processing Systems* 26 (2013).

Padilla, Thomas. “Responsible Operations: Data Science, Machine Learning, and AI in Libraries.” Dublin, OH: OCLC Research, 2019. https://doi.org/10.25333/xk7z-9g97.

Qingqing, Zhou, and Zhang Chengzhi. “Measuring Book Impact via Content-Level Academic Review Mining.” *The Electronic Library* 38, no. 1 (2020): 138–54. https://doi.org/10.1108/EL-08-2019-0184.

Řehůřek, Radim, and Petr Sojka. “Software Framework for Topic Modelling with Large Corpora.” In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50. Valletta, Malta: ELRA, 2010.

Steele, Tom, and Nicole Sump-crethar. “Metadata for Electronic Theses and Dissertations: A Survey of Institutional Repositories.” *Journal of Library Metadata* 16, no. 1 (2016): 53–68. https://doi.org/10.1080/19386389.2016.1161462.

Strader, Cynthia R. “Author-Assigned Keywords versus Library of Congress Subject Headings.” *Library Resources and Technical Services* 53, no. 4 (2011): 243–51. https://doi.org/10.5860/lrts.53n4.243.

Tait, Terry T. “Enduring Failures: A Borderlands History of the Iraq War and Its Aftermath.” MA Thesis, Miami University, 2019.

Tuarob, Suppawong, Line C. Pouchard, and C. Lee Giles. “Automatic Tag Recommendation for Metadata Annotation Using Probabilistic Topic Modeling.” In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*, 239–48. JCDL ’13. New York: Association for Computing Machinery, 2013. https://doi.org/10.1145/2467696.2467706.

Wiersma, Gabrielle, and Esta Tovstiadi. “Inconsistencies Between Academic E-Book Platforms: A Comparison of Metadata and Search Results.” *portal: Libraries and the Academy* 17, no. 3 (2017): 617–48.

Yelton, Andromeda. “HAMLET: Neural-Net-Powered Prototypes for Library Discovery.” In *Artificial Intelligence and Machine Learning in Libraries*, edited by Jason Griffey, 10–15. Chicago: ALA TechSource, 2019.

Zollers, Alla. “Emerging Motivations for Tagging: Expression, Performance, and Activism.” In *WWW 2007: Proceedings of the 16th International Conference on World Wide Web, Banff, (Alberta, Canada), May 8-12, 2007*, edited by Carrie Williamson and Mary E. Zurko. New York: ACM Press, 2007. http://wwwconference.org/www2007/workshops/paper_55.pdf.

Chapter 9

"We Could Program a 'Bot' to Do That!":

Robotic Process Automation in Metadata Curation and Scholarship Discoverability

Anna Milholland and Mike Maddalena

Introduction

Existing at the intersection of knowledge creation, aggregation, access, and preservation, academic libraries are well-positioned to collect and promote scholarship. Indeed, libraries do this in a myriad of ways, from gathering metadata and standardizing information discovery to preserving and ensuring open access, and to cutting-edge research using their institutional repositories. Yet, academic libraries and their parent institutions can also find themselves hampered in the quest to effectively capture research production and impact when information is dispersed across multiple, disparate knowledge bases.

As part of both the Raymond A. Mason School of Business and William & Mary Libraries, the McLeod Business Library has been an advocate for collecting, connecting, and promoting business faculty scholarship. The business librarian serves in an ex-officio capacity on the Faculty Research Committee (FRC), which is charged by the School of Business's Faculty Affairs Committee with promoting research and its impact. Over the past two years, the work of the FRC has primarily focused on creating a landing page that would showcase faculty scholarship and promote its discoverability while also connecting external scholars and members of the media with the School of Business's topical experts. Recognizing the value that information curation, organization, and management could bring to this process, the business librarian looked to campus partners to identify opportunities for promoting this objective.

Within the School of Business, the business librarian engaged the Office of Academic Affairs, which administers Digital Measures, the university's annual reporting system for teaching, scholarship, and service. Not only is Digital Measures useful for university reporting, but it also serves as a repository to support the School of Business's AACSB accreditation. To gain access to Digital Measures and run reports, the business librarian coordinated with the Office of Institutional Accreditation and Effectiveness, enabling access to the School of Business's Digital Measures profile.

William & Mary Libraries was an optimal and willing partner. Swem Library, the main university library, has long advocated for open access, as they manage and promote scholarship through W&M ScholarWorks, the university's open access repository. Additionally, Swem Library was interested in tracking and analyzing research impact and accomplished this using 1Figr data. They were simultaneously pursuing an ORCID ID campus rollout, and all parties had aspirations that both ORCID profiles and ScholarWorks profiles could be populated with Digital Measures metadata.

In essence, all parties hoped to synchronize data about scholarship, a vision that scholarly communication experts such as Liz Allen, Gabriela Mejias, Phill Jones, and Karin Wulf identify as a "sustainable research infrastructure."[1] While each differs in their exact articulation of this vision, they share the notion that "a strong and sustainable research infrastructure—the tools, services, and systems that support the research process—is vital. It speeds up the dissemination of research, helps ensure that it's FAIR (findable, accessible, interoperable, reusable), minimizes the risk of errors, and reduces the administrative burden on researchers and their organizations alike."[2] In essence, research information can be accessed and communicated for optimal discovery across multiple platforms.

Harvesting information from Digital Measures proved challenging because much of the data was entered manually, rather than being sourced from indexes like CrossRef and ORCID. Not only were major fields—including DOIs, URLs, ISBNs, and ISSNs—left blank, but much of the other bibliographic data—including journal titles, publication dates, volume, and issue information—was not standardized. While this challenge resulted in an opportunity to scrub Digital Measures metadata and fill in the gaps, doing so became time-consuming. As a response to this, the business librarian looked for ways to automate data retrieval and populate Digital Measures using retrieved metadata.

Automation and RPA in Academic Libraries

Attributed to D. S. Harder of Ford Motor Company, the term "automation" has been in existence since the mid-1940s,[3] though using machines to replace human labor predates the term itself. The first evidence for this is documented by Ctesibius (285–222 BCE), who applied the first known feedback control system to a water clock.[4] By replacing human labor with machines, the need for automation eclipsed factories and became ingrained into a number of industries and organizations, including libraries, as they began to rely on computers and other forms of artificial intelligence, including robots, to work faster and smarter by improving processes.

The visionary Alan Turing wrote, "I propose to consider the question, 'Can machines think?'"[5] As public-facing entities with limited staffing and budgets, libraries often hope that elements of automation and artificial intelligence can think enough—or as well as programmers' program them to think—to serve as adequate replacements for human staff, who can be redeployed to support other mission-critical services. But how do libraries determine the best candidates for automation?

"Since the 1960s, libraries have used technology in general, and computers in particular, to automate a wide range of administrative, public, and technical services tasks,"[6] including descriptive cataloging. As original descriptive cataloging can be labor-intensive, copy cataloging has been a viable, cost-effective alternative since the 1960s, with the advent of Machine-Readable Cataloging (MARC) records. Among other organizations, OCLC has long been a repository for these machine-readable cataloging records,[7] with WorldCat, its union catalog, a library staff, and a public-facing portal for discovery of local and global library holdings. Through the Z39.50 protocol, these MARC records can be accessed and imported seamlessly into an ILS, thus providing catalogers with high-quality records and reducing time spent on original cataloging. The Z39.50 also supports operations such as Interlibrary Loan, where staff cannot only manually search other libraries' holdings through WorldCat but they can also use WorldCat and its rich bibliographic data—including OCLC numbers, DOIs, and ISBNs—as a conduit for initiating both mediated and unmediated borrowing requests from one library to another.

These backend library technologies are classified as heavyweight information technology, or heavyweight IT—that is, "a knowledge regime, driven by IT professionals, enabled by systematic specification and proven digital technology, and realized through software engineering."[8] While heavyweight IT provides necessary infrastructural and automation support, it isn't necessarily agile or conducive to innovation. Contrast this with lightweight IT, which is defined as "a knowledge regime, driven by competent users' need for solutions, enabled by the consumerization of digital technology and realized through innovation processes."[9] Lightweight IT, as a rule, is accessible by general users and may be deployed by those users to solve problems.

Robotic process automation (RPA) software, a lightweight IT solution, enables users to construct and deploy robots via a graphical user interface (GUI), rather than a command line. This software theoretically makes bots, which are designed to emulate basic human tasks, accessible to programmers and non-programmers alike. RPA is used across industries, particularly in business to automate tasks, but its usage in libraries remains unexplored until now. How, then, should information professionals apply relevant software like RPA—and bots—for information retrieval and production?

Choosing RPA Software

Automation—and RPA—can be attended or unattended, according to Gartner, a leader in and producer of technology-focused market research. Attended RPA focuses on the end-user, as RPA tools can "extract information from systems and related documents, shaping it and preparing it for consumption."[10] Unattended RPA, on the other hand, deploys scripts to accomplish rote tasks that an individual would typically complete one at a time. Considering the amount of time that would be involved in searching article titles individually, retrieving and documenting each DOI to a spreadsheet (or even copying to the clipboard), and then adding those individually to Digital Measures, Business Intelligence and the Business Library determined that the process was high-volume, non-complex, and time-consuming. As such, it could be programmed as a repetitive, unattended RPA practice.

According to Gartner, "Leaders in a market combine an insightful understanding of the realities of the market, a reliable record, the ability to influence the market's direction, and the capability to attract and keep a following."[11] Its Magic Quadrant report depicts three RPA leaders: UiPath, Blue Prism, and Automation Anywhere. Yet, a company's leader designation is not necessarily an indication that their specific RPA solution will be in the customer's best interests.

Since 2019, William & Mary undergraduate business students have had access to RPA software and bots, courtesy of a gift from UiPath, which was named the fastest-growing North American technology company in Deloitte's Tech Fast 500 ranking.[12] UiPath offers a suite of RPA services, including products that are tailored to both business users and developers to automate tasks and manage robots and processes used across industries. This gift also supports a three-year partnership between William & Mary and UiPath, and it serves to broaden the integration of RPA technology on campus. Given this unique relationship with UiPath, the need for a low-cost solution, and the desire to innovate using lightweight IT, the authors opted to use UiPath's RPA software to automate this process.

Process Automation Methodology

A Center of Excellence (CoE) is a way to embed RPA effectively into the organization and to redistribute accumulated knowledge and resources across future deployments.[13] The use of a CoE is necessary for a full fleet of complex calculations, and the Raymond A. Mason School of Business's Business Intelligence (BI) team doubles as a small center of excellence. Since the Business School has a small CoE, the focus is to automate processes that meet specific characteristics that will be relatively easy to maintain and have a high impact and/or return on investment. When this metadata project idea was first brought to the BI team, it immediately fit as a potential candidate that matches many of the characteristics the team looks for when automating a process. The BI team leans on a process complexity matrix developed for the Business School's "RPA for Business" class to determine what processes could be a good candidate for RPA, and BI uses some of the criteria from the automation matrix for rapid development:

- process is repetitive
- decisions are rule based
- limited human judgment required
- process inputs and outputs are structured
- the system environment does not constantly change

Before developing the process programmatically, the authors compartmentalized the process to allow for easier development. The authors broke down the process into five stages: creating the input file from Digital Measures, reading the input file, sending data to WorldCat, retrieving data from the online catalog, and writing to the output file. From this, the team developed a requirements document that outlined goals for each process category, what was developed, and what was needed to be achieved after development. Development consisted of reading the locally held Excel input file of Digital Measures faculty publications; taking the article title and searching in WorldCat, thus retuning a search hit; opening up the search item and data scraping the DOI before writing it back to the initial Excel document; looping back with the next item on the list and repeating until reaching the end of the list. The BI team used UiPath’s Development Studio to develop each of the process categories, and the team leveraged some of UiPath’s recording technology (which allows for recording steps, similar to a macro) along with .NET frameworks (Microsoft’s developer network containing tools, programming, and libraries) to program the process automation or less formally “bot.” The Development Studio, which is a graphical user interface, enables workflow visualization, including tree charts or screenshot-type workflow steps, which can make processes more easily translated from stakeholder to developer. It is generally more user-friendly for non-programmers to understand (and a visual of UiPath studio is attached in the appendix). While developing the “bot,” the team used UiPath’s version control platform as well as Github to track major and minor revisions to the process.

Once initial development was completed, the authors conducted *beta* testing to see how the process would perform. After the first test, it was apparent that the authors needed to revisit and revise several areas of process development. One of the problematic areas involved leveraging the search functionality in the WorldCat catalog. Unless the bot searched for the exact published title, WorldCat wouldn’t necessarily return the correct research article. In order to solve this issue, the BI team added quotes to the search criteria to enforce an exact match. Doing this obviously increased the accuracy of identifying the research, but it also reduced the number of research article matches because of the exactness of the title required. This is not a perfect solution and requires staff to research the entries with no results returned, as the title may be slightly different. The second problem the authors encountered was that the selected identifiers needed to be tuned so that they would always select the correct button to pull up the DOI. The authors created wildcards to represent “begins with” when selecting the HTML identifiers specifically in div (content divider) objects to solve this issue. After facing these challenges, the authors were ready for production as an unsupervised automation or an automation that doesn’t have to be monitored as it runs.

Measuring Results and Impact

When measuring the impact of process automations, it is important to know how many resources were saved leveraging this technology to substantiate the business case for transforming this process. This measurement also serves as a process improvement. There are several ways to measure the impact or the return on investment of automating processes. One way is to measure the time it takes one person to look up one item from the spreadsheet, then multiply it by the number of items and set that as the time baseline of the process. RPA users can then measure how long it takes the "bot" to look up all of the items and analyze the performance versus the baseline; in finance, this productivity about the baseline would be referred to as *alpha*.

Another way to measure the results is by taking the manual process and identifying the accuracy and comparing that with the accuracy of the automated process. In both cases, RPA users hope to see positive alpha in both efficiency and accurateness. If not, RPA users must decide as there might be tradeoffs in accurateness for efficiency, which may or may not be acceptable depending on the process or situation. There are also qualitative impacts that may be obtained via survey, including but not limited to the following: the job satisfaction of the people performing the task, before and after the process was automated; the work-life balance before and after a process has been automated; and the workplace culture now that a part of an individual's job has been automated.

The BI team includes qualitative measures when the routine that is automated impacts someone's job more than several times per year or is an invasive organizational process. The team defines this as a process that someone's core job or jobs are based on, at minimum, once a week. Generally speaking, qualitative results for sporadic automation have limited or no impact on a person's overall job and are muted in the survey results. For this automation, the authors only measured the quantitative measures, as it is only carried out several times a year. This automation had 3,671 items to look up and return a result, which yielded 766 DOIs from WorldCat. When a library staff member searched after the bot, they yielded 843 DOIs. While at first glance the human might appear more productive, the "bot" can check the system in eight seconds whereas a person takes on average about thirty seconds, which is 73 percent faster than a human. The entire process took the "bot" 8.16 hours and employees a cumulative 30.59 hours to complete.

When analyzing the time performance of the "bot," the authors determined that it produced a decent return, especially when considering this process occurs twice per year. Though the bot's performance isn't perfect due to the search result constraints and employees took longer (because they had to comb through the results to find the correct match), this comparison is not perfect, as they aren't *ceteris paribus* (all else equal). With that said, the "bot" is successful at retrieving the low-hanging fruit, even though more ambiguous records require manual research, thus yielding a complementary automation.

Analysis of Automation

This process was hamstrung by the limited performance of the WorldCat search, with its inability to provide accurate results without a direct quote search. The lesson learned from this is that an automation is only as good as the systems being used. Systems can be a major limitation to fully automate a process, and that has to be accounted for when approaching an automation project. The computers being used and the quality of the internet connection also have a direct impact on the speed of the automation.

The authors used a performance computer with an i7 processor and 16 GB of RAM and a 100 MBPS internet connection allowing the automation to move quickly without many issues. If this automation was run with an i3 processor and 4 GB of RAM and a 5 MBPS internet connection, the speed would be reduced, an obstacle RPA itself cannot overcome. Automation speed is only as good as the technological infrastructure within which the team can operate.

Another limitation with RPA is the lack of built-in or turnkey artificial intelligence (AI). The inexact search results proved a limitation. One way to potentially solve this would be to allow the search to run and have AI provide the best match based on the numerous results. While the BI team could have included a separate Python-based machine learning model to give the best match, the long-term maintenance and sustainability of the automation would have made it far too complex to maintain. Frequent structural or systematic website, system, or application changes negatively impact the ability to automate efficiently. As AI packages become more common and robust, the authors believe there will be more turnkey AI solutions available that alleviate these issues.

Currently, RPA as a technology uses selections in the code to look for specific identifiers; if the item moves and the identifier remains the same, RPA can still identify the item. However, if the codes change in the background and RPA loses the identifier, this generates an error. The authors also expect as time passes that RPA will include that type of technology, which allows for more citizen developers, as it will almost be a point-and-click development environment. This will also decrease the labor intensiveness of a CoE as the maintenance would be reduced.

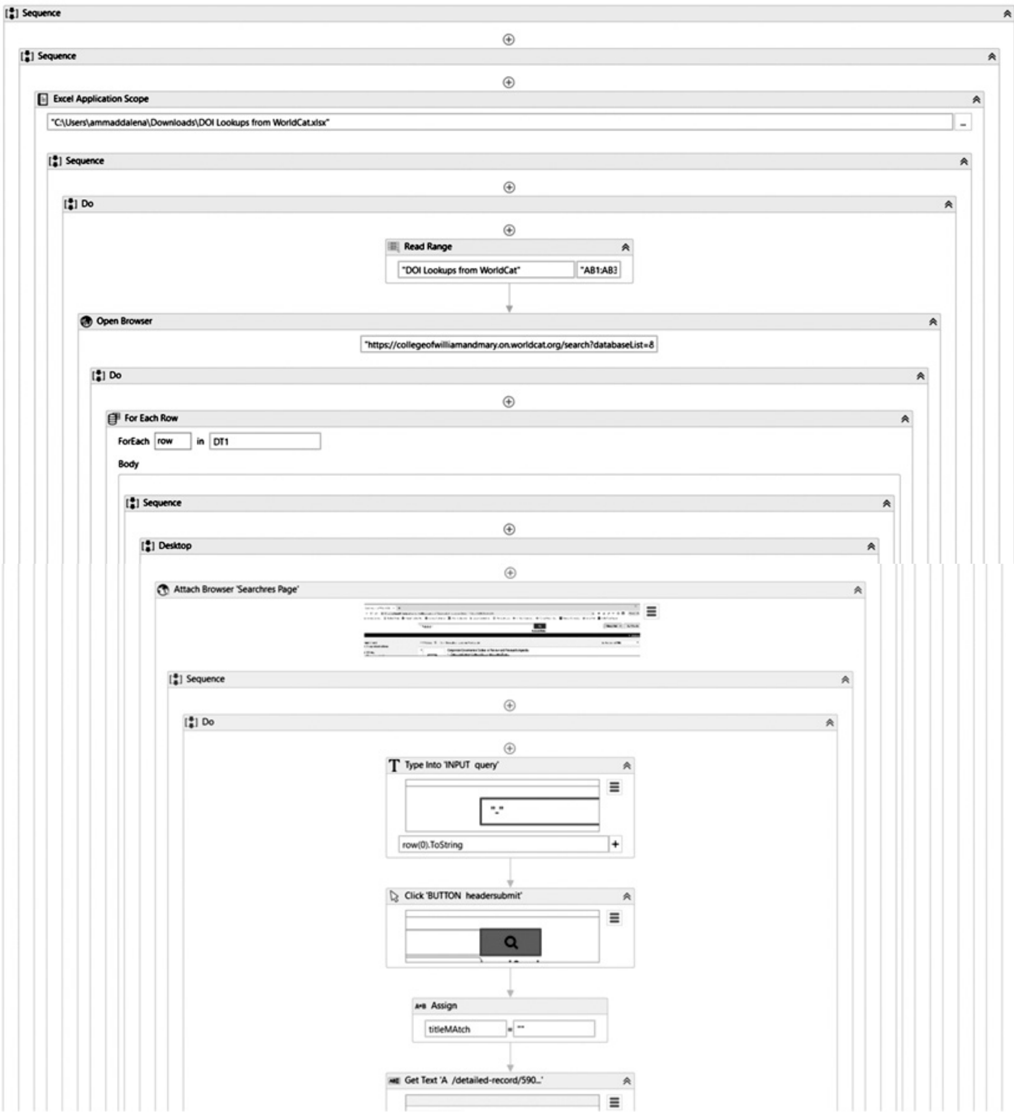
Conclusion

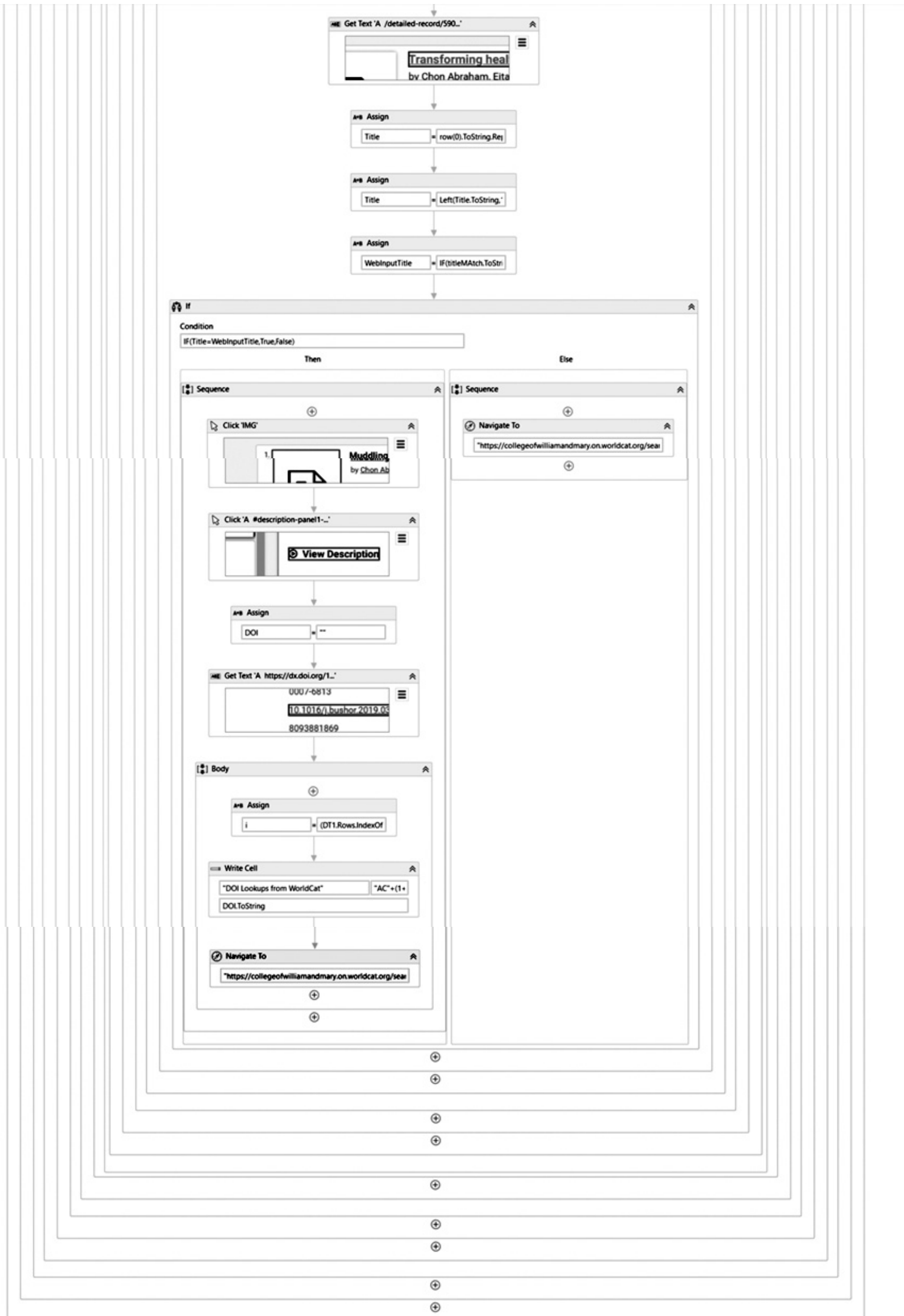
As with all initiatives, it is important to note the lessons learned and modify practices as part of project assessment. Business Intelligence retains all requirement documents, key takeaways, lessons learned, and code in an RPA repository to reference for maintenance purposes and future development, including retrieving ISBNs and ISSNs for more complete bibliographic data to include within Digital Measures. As the Digital Measures publication data becomes more robust and more reliable with this information, we hope to leverage that data to discover faculty open access publications using SherpaRomeo and related tools. Depending on a journal's open access and copyright archiving policy, qualifying content could be batched into W&M ScholarWorks Institutional Repository and the Expert Gallery and could then be shared with the masses. In combination with other licensed evaluation tools, including Scopus, SciVal, and ORCID, this use of Scholar-Works could enable the Raymond A. Mason School of Business to track faculty members' research impact across platforms, highlight core research strengths, and promote research expertise.

Both McLeod Business Library and Business Intelligence will continue to share data for process improvement, particularly as we develop a dashboard of metrics collected from the data. Interactive dashboards provide useful insight, as well as data-driven decision-making, which are conducive for distilling and displaying results. These dashboards, once developed, can be integrated into the automation itself. The authors also expect to revisit this, as the process has been ongoing, to identify any improvements that could be made to the automation. Though this isn't necessarily an easy or a perfect solution, it still provides value over the baseline and is considered to be successful automation. As this was the first known and successful usage of RPA—and UiPath specifically—by a Business Library and Business Intelligence unit and the first usage of RPA within the Raymond A. Mason School of Business, the authors anticipate that this work could present additional opportunities for collaboration with the libraries and the university.

Appendix

Figure 9.1




Get Text 'A /detailed-record/590...'
Transforming heal
by Chon Abraham, Eita
Assign
Title
row(0).ToString.Re
Assign
Title
Left(Title.ToString,'
Assign
WebInputTitle
IF(titleMAtch.ToStri
If
Condition
IF(Title=WebInputTitle,True,False)
Then
Else
Sequence
Sequence
Click 'IMG'
Muddling
by Chon Ab
Navigate To
"https://collegeofwilliamandmary.on.worldcat.org/sea
Click 'A #description-panel1-...'
View Description
Assign
DOI
""
Get Text 'A https://dx.doi.org/1...'
0007-6813
10.1016/j.bushor.2019.0
8093881869
Body
Assign
i
(DT1.Rows.IndexOf
Write Cell
"DOI Lookups from WorldCat"
"AC"+(1+
DOI.ToString
Navigate To
"https://collegeofwilliamandmary.on.worldcat.org/sea

Endnotes

1. Alice Meadows, "Building a Sustainable Research Infrastructure," *The Scholarly Kitchen*, April 30, 2020, https://scholarlykitchen.sspnet.org/2020/04/30/building-a-sustainable-research-infrastructure/.
2. Meadows, "Building."
3. "Automation," Encyclopedia Brittanica, accessed May 1, 2020, https://www.brittanica.com/technology/automation.
4. Massimo Guarnieri, "The Roots of Automation Before Mechatronics [Historical]," *IEEE Industrial Electronics Magazine* 4, no. 2 (June 2010): 42.
5. Alan Turing, "I. Computing Machinery and Intelligence," *Mind* 59, no. 236 (October 1, 1950): 433.
6. William Saffady, "Library Automation: An Overview," *Library Trend*s 37, no. 3 (Winter 1989): 269.
7. Saffady, "Library Automation," 274
8. Bendik Bygstad, "Generative Innovation: A Comparison of Lightweight and Heavyweight IT," *Journal of Information Technology* 32, no. 2 (June 2017): 191.
9. Bygstad, "Generative Innovation," 191.
10. Derek Miers, Marc Kerremans, Saikat Ray, and Cathy Tornbohm, "Magic Quadrant for Robotic Process Automation Software," Gartner, July 8, 2019.
11. Miers, Kerremans, Ray, and Tornbohm, "Magic Quadrant."
12. "UiPath Ranks No. 1 Fastest Growing Company in North America on Deloitte's 2019 Technology Fast 500," *Bloomberg* (November 7, 2019). https://www.bloomberg.com/press-releases/2019-11-07/uipath-ranks-no-1-fastest-growing-company-in-north-america-on-deloitte-s-2019-technology-fast-500.
13. Mandi Bishop, Cathy Tornbohm, Derek Miers, "Scale Automation in Healthcare Using a Center of Excellence," Gartner, January 28, 2020.

Bibliography

Bishop, Mandi, Cathy Tornbohm, and Derek Miers. "Scale Automation in Healthcare Using a Center of Excellence." Gartner. January 28, 2020.

Bloomberg. "UiPath Ranks No. 1 Fastest Growing Company in North America on Deloitte's 2019 Technology Fast 500." November 7, 2019. https://www.bloomberg.com/press-releases/2019-11-07/uipath-ranks-no-1-fastest-growing-company-in-north-america-on-deloitte-s-2019-technology-fast-500.

Bygstad, Bendik. "Generative Innovation: A Comparison of Lightweight and Heavyweight IT." *Journal of Information Technology; London* 32, no. 2 (June 2017): 180–93. http://dx.doi.org.proxy.wm.edu/10.1057/jit.2016.15.

Encyclopedia Brittanica. "Automation." Accessed May 1, 2020. https://www.britannica.com/technology/automation.

Guarnieri, Massimo. "The Roots of Automation Before Mechatronics [Historical]." *IEEE Industrial Electronics Magazine* 4, no. 2 (June 2010): 42–43. https://doi.org/10.1109/MIE.2010.936772.

Meadows, Alice. "Building a Sustainable Research Infrastructure." *The Scholarly Kitchen*. April 30, 2020. https://scholarlykitchen.sspnet.org/2020/04/30/building-a-sustainable-research-infrastructure/.

Miers, Derek, Marc Kerremans, Saikat Ray, and Cathy Tornbohm. "Magic Quadrant for Robotic Process Automation Software." Gartner. July 8, 2019.

Saffady, William. "Library Automation: An Overview." *Library Trends* 37, no. 3 (Winter 1989): 269–81.

Turing, Alan M. "I.—Computing Machinery and Intelligence." *Mind* LIX, no. 236 (October 1, 1950): 433–60. https://doi.org/10.1093/mind/LIX.236.433.

Chapter 10

More Than Just Algorithms:

A Machine Learning Club for Information Specialists

Mark Bell and Leontien Talboom

Introduction

Over the last few years, artificial intelligence (AI) and especially machine learning (ML) have become increasingly prominent in most industries, with the humanities sector being no exception. A number of institutions across the galleries, libraries, archives, and museums (GLAM) sector have been experimenting with and implementing algorithms or other computational techniques. Examples include The Living with Machines project at the British Library and the Alan Turing Institute, and The Museums + AI Network.[1] Specialised academic labs, like Yale's DHLab and Oxford's Visual Geometry Group, are also doing work related to ML.[2] At The National Archives (TNA) of the United Kingdom, activities include co-organising both the Computational Archival Science Network and an explainable AI workshop[3] as well as hosting an AI symposium for the archive sector and last year's Annual Digital Lecture on the topic of algorithmic bias.[4] TNA's Digital Strategy includes applying ML for appraisal, selection, and sensitivity review as well as improving access to the collections. It also emphasises the importance of developing "digital capability, skills and culture" within the organisation.[5]

Literature Review

A recent Archives, Access and AI conference showcased a number of projects that either use AI or aspire to do so.[6] The conference was not limited to speakers from the archives, as the title would suggest, but from organisations all across the humanities sector. For this chapter, the term *information specialist* will be used to refer to people from across this sector who acquire, appraise, and preserve materials. This group is starting to understand that their role with regard to computational methods, including ML, is changing. Public awareness of ML tools is also increasing. A 2016 survey by Ipsos Mori and the Royal Society found that only 9 percent of those surveyed recognised the term *machine learning*, yet 76 percent of respondents were aware of applications such as speech recognition and question answering, even if they didn't know they were powered by ML.[7] A similar poll in 2019 conducted by Kantar Public found that only 7 percent of respondents had never heard of AI, while 12 percent thought they knew a lot.[8] It should be noted that the term AI is far more ubiquitous in popular culture than ML, which may explain the disparity. Increasing awareness is partially explained by the increased automation in people's lives but also due to initiatives such as the Royal Society's Machine Learning project and the Finnish Government's Elements of AI course, which aims to educate 1 percent of European citizens on the basics of AI by 2021.[9]

There is also a growing societal awareness of what data can and cannot be used for, along with increased recognition of what can happen without the correct safeguards in place, such as amplifying existing biases within the data.[10] Mordell warns that all the social justice work done across the sector could be undone by the implementation of computational methods.[11] Griffey also warns about similar implications if these tools are not approached with caution.[12] Johnsson, Jakeway, et al. talk about how this technology is not only technical but also social and far more subjective than may be thought, as it relies on human judgment and biases.[13] It is not only about being cautious when implementing these tools; a number of papers have highlighted how critical the information specialist's skills—such as appraisal, selection, and cataloguing—are in the digital age.[14] Some researchers, however, argue that highlighting the problems and benefits is not enough and that there needs to be an emphasis on how important it is to engage in these discussions.[15]

Information specialists face other barriers when they implement these tools and experiment with them. Common barriers to AI projects include having insufficient data in the right format and insufficient skilled resources to take experimentation forward. As a result, growth can be witnessed in automated ML products, such as Google's AutoML,[16] which aim to democratise the building of models. The rationale is simple: data scientists are rare and therefore expensive and difficult to recruit, whereas there are millions of software developers already embedded into organisations.[17]

While processing a dataset of labelled example records (training data) through a proprietary "black box" algorithm is generally cheaper than hiring a data scientist, there is a loss of control over the process, including the ability to adjust the results and to explain the methods. Automated approaches take the focus away from the algorithms, which are hidden, and put it back firmly on the data that is used to train them. Information specialists will become critical to the selection and creation of training data. This is a paradigm shift from a world where software developers elicit rules from users then design and develop a system that implements those rules. Subject matter experts now need to communicate with data scientists about selecting the right model that suits both the data and the application. These decisions are often a balance between accuracy and explainability.

Explainability is a rising trend in the debate around AI. Machine learning can be separated between statistical and algorithmic approaches, which Breiman describes as the "two cultures."[18] While both ultimately result in predictions, the statistical approach begins with identifying underlying models that describe physical phenomena, whereas algorithmic approaches are results-focused. Deep learning is used for complex tasks, such as image recognition and handwritten text recognition.[19] The "deep" part refers to the depth, or number of layers, in a neural network algorithm, each layer being a matrix of weights that are applied to the input data as it passes from layer to layer. The deeper the network, the more generalisable it becomes, but depth leads to greater complexity.[20]

Machine learning algorithms are evaluated against benchmark datasets, often termed the common task framework (CTF).[21] While this has led to incredible progress, many computer scientists are more focused on the performance of the tool than the understanding of how it functions,[22] which has led to some of the leading researchers in the field referring to it as "alchemy."[23] There have been attempts made to better understand neural networks with projects such as The AI Detectives.[24] Efron and Hastie consider empirical approaches like the CTF to be "ultimately unsatisfying without some form of principled justification."[25] While they are optimistic that the statistical inference community will eventually connect modern machine learning algorithms to a "central core of well-understood methodology," the issue remains that highly complicated tools are being built and the understanding of their internal workings is limited.[26] There is also the added problem that the benchmark datasets are not representative of the collections that information specialists would like to process with algorithms.

In order to address the growing interest in AI at TNA and the growing concerns around the ethical implications of these tools, a set of workshops entitled Machine Learning Club was organised. Members from all areas of the organisation attended the sessions. The aim of these workshops was not to turn information specialists into data scientists but rather to develop an understanding of what ML can offer to archives and which skills are needed to make the implementation of these tools successful. The sessions were designed to prepare staff to identify opportunities, remain alert to pitfalls, and be able to engage confidently with these exciting new technologies. This chapter explains how the Machine Learning Club was created and the content covered in the session. The knowledge and confidence gained from participating in the club are also discussed.

Machine Learning Club

The Machine Learning Club (MLC) was created to respond to a growing interest in the digital preservation team at TNA for more hands-on experience with ML. In order to pilot this initiative and measure the level of interest across the organisation, the authors organised a series of lunchtime talks. These talks focused on different aspects of ML, starting with data preparation and a discussion of some well-known algorithms. Each talk lasted an hour and included further readings and examples of where to get practical experience. The club was well attended with around thirty attendees each session. Participants received homework assignments every session; however, there was little engagement with these assignments. Feedback received during the concluding session included a desire from participants to get more hands-on experience with this technology.

The authors initially selected Machine Learning Mastery and Towards Data Science as examples of online ML tutorials for the lunchtime talks.[27] These tutorials, however, can be highly technical and geared toward budding data scientists as well as require specific computer applications that present logistical challenges with regard to installation. The authors decided to modify the lunchtime talk sessions and create meaningful and relatable tutorials for information specialists. They presented these in four three-hour workshops offered monthly, providing aid and guidance when needed. Due to the large time commitment and the incremental nature of the workshops, the participants had to register and confirm with their manager that this time could be spent on MLC.

The organisers chose to use Google Colab as the environment to run the tutorials to avoid the technical issues surrounding installation of software applications and the lack of technical skills.[28] The Google-hosted "notebook" environments are run from the browser, making it unnecessary to install any software, and the computer code is run by clicking a play button in a cell.[29] The participants could therefore focus on the results given by the code and not on trying to get it to function. The tutorials and data were hosted on Google Drive, an environment familiar to most, which removed the barrier of trying something new. An example of one of the Google Colab tutorials from the MLC can be seen in figure 10.1.

Figure 10.1

Part of an MLC tutorial, which is hosted on Google Colab. The first part of the tutorial consists of a piece of code that can be run by clicking the play button on the left-hand side. This will run the code and produce the underlying graphs as an output.

Session 1

The first MLC session focused on data analysis and data cleaning, which are important first steps in an ML workflow. The organisers started the session with a survey to measure the participants' knowledge of ML and to understand their motivations to take part in these workshops. They also used this information to modify the course materials for subsequent sessions. The presenters provided a general overview of the course and explained some basic ML concepts. The participants then took part in three tutorials, which involved different techniques and increasing levels of complexity.

The organisers selected a simple zoological dataset for the first tutorial.[30] Given that the dataset was in good condition, the presenters added a number of mistakes to determine if the participants could identify them. The tutorial focused on simple visualisation techniques, such as bar charts, for exploring numerical datasets. The second tutorial explored a US Census Income dataset, which was both numerical and categorical, making it more complex than the first dataset.[31] The participants learned statistical summarisation techniques, such as medians, averages, and quartiles, to aid them in understanding the data. This dataset also presented opportunities to discuss biases within the data. The final dataset, an Amazon food reviews dataset, was predominantly textual, which meant introducing word clouds and other text-mining techniques to be able to visualise this data, as the statistical techniques discussed in the other two tutorials would not be adequate for textual analysis.[32]

During this session, the presenters taught a basic understanding of the methods that can be used for different datasets and focused on the importance of data cleaning and gaining a thorough understanding of the data before using any type of ML. In order to encourage the participants to reflect on the activities and encourage wider group discussions, the code in the Google Colab notebooks was interspersed with questions.

Session 2

During the second session, the organisers provided the first hands-on experience of two ML algorithms: *nearest neighbours* and *decision trees*.[33] The participants began the workshop with a number of paper-based games (figure 10.2) to help build some intuition about how ML algorithms function. *Nearest neighbours* was introduced via a grid of playing cards, indexed by suit and number, and the participants had to decide whether to put down a blue- or red-backed card depending on the colour of nearby cards. To understand *decision trees*, the group was split into two and majority voting was used to build a tree for deciding whether to go outside depending on what the weather was. A third activity was used to explain the concept of *decision boundaries*,[34] where participants learned that an algorithm will change from predicting one class to another. Teams had to place pieces of strings to divide the squares and triangles plotted on paper; there was a penalty system in place (e.g., a shape being on the wrong side of the string) that replicated the optimisation process used by ML.

Figure 10.2
String, playing cards, and coffee: hands-on machine learning.

Following these exercises, there was a brief presentation to put them in context. The organisers then presented a number of tutorials to provide hands-on experience of the algorithms. In addition to the activities with the algorithms, the presenters explained two ways of assessing their performance: accuracy scores and confusion matrices. When participants needed more support, the presenters reviewed content from the previous session. The datasets used during this session were the same as the ones used in Session 1.

Sessions 3 and 4

Due to the COVID-19 pandemic lockdown, the organisers had to adapt the workshops to a new learning environment. They had originally planned to introduce an additional set of algorithms and compare them to the algorithms from Session 2. The organisers decided to keep this approach and include more content since they were no longer limited by the time restrictions of a classroom session. For this session, one data source was used: a set of categorised records from the Discovery catalogue at TNA.[35] The organisers chose this data source mainly because it is an internal dataset that is both familiar and relevant to most of the participants and because it is a prime candidate for future applications of ML. They created four tutorials that exemplified a small ML project, with each tutorial focusing on a specific aspect of the ML process: data analysis, data preparation, ML classification, and interpretation/explanation. For each tutorial, participants had to use data introduced in the previous section, meaning that data selection decisions made in the first part of the tutorial could influence the ML accuracy in the third part. The goal was to demonstrate that choices made throughout the ML workflow could have an impact later in the process.

Session 3 is currently the last session of the ML Club, and the organisers hope to hold a fourth one once the situation around COVID-19 is resolved. They aim to focus the last session on the explainability of ML tools. In the meantime, the organisers have decided to postpone this session until it can be run in a classroom environment. They intend for the session to be heavily discussion-based, and their aim is to gather perspectives informed by the attendees' differing backgrounds. The intention is to break new ground in exploring the potential of ML from an information specialist viewpoint.

Discussion

Given the fact that these workshops were a new initiative, the organisers began Session 1 with a survey to gain a better understanding of the skills of the information specialists and their expectations with regard to the workshops. By combining the survey results with the participants' feedback during the sessions, the presenters were able to design each workshop around the needs of the attendees.

The organisers' decision to use Google Colab as the main environment was very beneficial. The environments could be run from any computer or device, without spending a lot of time downloading and installing software, which left more time for teaching and discussions. The information specialists were able to engage in the hands-on experience they were hoping for due to Google Colab. The attendees with no Python experience could still run code and interpret the results, while experienced coders could delve into the functions being used. Additionally, individuals with limited knowledge of Python attempted to experiment with the code, which was a positive and surprising outcome.

While the simple design and code initiation of Google Colab was a benefit, it was also its greatest drawback. Participants can simply press the activation button and not fully engage with the material. To solve this issue, the organisers included regular questions in the tutorials to prompt a more critical analysis of the visualisations or numerical summaries, and participants were also encouraged to work in groups and discuss amongst themselves. Group discussions had little success with most participants working in silence; however, participants were not afraid to ask questions. A participant in Session 1 asked, "What do you mean by algorithm?" which led to a discussion about the difference between standard algorithms being more analogous to recipes and the ML algorithms, which detect and learn from patterns in data.

To encourage group discussion and better exemplify ML algorithms, the organisers made changes to their approach for the second session. They created paper exercises and designed activities that required teamwork to test a wide range of parameter settings. Discussion was an important element of the classes and led to a number of interesting questions that showcased how the different perspectives of information specialists can be very beneficial to the ML conversation. They were particularly interested in data provenance, and their questions often highlighted the importance of contextualisation to information specialists, which would not necessarily be of concern to computer scientists.

The organisers noticed that the participants needed to understand everything, which led to several participants engaging in deep research about zoological classifications and causing some consternation when presented with an algorithm that was too complex to explain.

The exercises in the second session gave participants an opportunity to "think like an algorithm," and the presenters frequently referred to them in the sessions to enable participants to relate what was happening in Google Colab with their experience of placing a piece of string between points. Most participants had no issues with the level of Session 1, but Session 2 appeared to be more challenging for many. This discrepancy is likely due to the fact that the tools of data analysis (graphs, numerical summaries, word clouds) are familiar, whereas ML introduces a number of new and often abstract concepts. A number of participants have therefore taken the opportunity to revise the older material before commencing the Session 3 tutorials.

A participant in Session 3 expressed that they could not conceptualize the rationale for using an ML approach when there was no improvement of the results compared to the current manual process. Although the question can be addressed in terms of this specific archival problem, it highlights participants' expectations of AI technologies and shows that the techniques presented in a tutorial are not necessarily those used for real-world applications. At the same time as these workshops, two participants were involved in an AI project with external suppliers. They mentioned that it was useful to be able to employ their new knowledge in discussions with data scientists and to relate supplier presentations of their ML workflows back to the course materials.

After completing the sessions, a final survey was sent to participants to evaluate their final opinion on the MLC. The comments were overwhelmingly positive, with all respondents finding the MLC useful. Not all respondents were able to apply the material directly to their day-to-day tasks, but a number did see the benefit of being able to understand the basic concepts: "Even if I might not be directly applying the learning, it means that when I hear others talking about the topic, especially in professional environments, I have a much greater appreciation of what they are discussing and why it is significant."

Furthermore, respondents emphasised the importance of understanding ML for information specialists: "It's a necessity. As an archive, we will need machine learning to help us carry out our responsibilities." Another respondent agreed with this: "I can see so many applications for it in archives that I would consider it essential for anyone working (or planning to) work with digital records."

The organisers of the MLC received permission from TNA to create the workshops, and the time and space to hold the workshops were kindly provided. Unfortunately, not all organisations will have this flexibility; therefore, the Google Colabs have been made available.[36] The code may help others who are creating similar workshops or could act as an inspiration. As mentioned previously, few online tutorials may be as useful to information specialists; however, some examples aimed at the humanities sector could also act as an inspiration, such as the Programming Historian, the GLAM Workbench, the Archives Unleashed Project, and the CLARIAH Media Suite.[37] Each employs similar platforms to showcase the use of computational methods in the humanities.

Conclusion

The authors designed the Machine Learning Club to help information specialists understand what can and cannot be accomplished with ML; however, the goal was not to train future data scientists. Hopefully, participants have gained confidence and knowledge on this topic, which will make it possible for them to participate in discussions surrounding the implementation of ML across the GLAM sector. Moreover, participants may be more willing to join the conversation on archiving these methods and techniques for future re-use.

Information specialists have skills that are relevant to the explainable AI debate but often lack the confidence to participate in these conversations due to their lack of knowledge of the underlying computational methods. The authors are hopeful that the basic knowledge shared in the workshops will motivate the information specialists at TNA to stay engaged with this field and its opportunities while understanding its drawbacks. Machine learning has great potential within the GLAM sector and information management more generally, but it is also important to understand how it will impact information specialists. The club has hopefully made it clear that algorithms are only one part of the process and that people who understand records and data are as important as ever. While interesting and effective, perhaps AI is not the solution to everything.

Endnotes

1. "Living with Machines," The Alan Turing Institute, 2020, https://www.turing.ac.uk/research/research-projects/living-machines; Oonagh Murphy and Elena Villaespesa, "The Museums + AI Network," 2020, https://themuseumsai.network/about/.
2. "Digital Humanities Lab," Yale University Library, 2020, https://dhlab.yale.edu/; "Visual Geometry Group," University of Oxford, Visual Geometry Group, 2020, https://www.robots.ox.ac.uk/~vgg/.
3. Eirini Goudarouli, "Computational Archival Science (CAS): Exploring Data, Investigating Methodologies—Workshop Invitation," *The National Archives Blog* (blog), 2019, https://blog.nationalarchives.gov.uk/computational-archival-science-cas-exploring-data-investigating-methodologies-workshop-invitation/; Jenny Bunn et al., "Workshop on Human-Centred Explainable Artificial Intelligence," *UCL Blog* (blog), 2019, https://blogs.ucl.ac.uk/hexai/.
4. "Symposium 'Archives and AI' by FAN and the UK National Archives," International Council on Archives, 2019, https://www.ica.org/en/symposium-archives-and-ai-by-fan-and-the-uk-national-archives; Safiya Umoja Noble, "Annual Digital Lecture: Algorithms of Oppression" (Annual Digital Lecture, The National Archives, Kew, 2 April 2019), YouTube, https://www.youtube.com/watch?v=n3dQUYTN9PA.
5. The National Archives, *Digital Strategy 2017–2019* (London: The National Archives, 2017).
6. Lise Jaillant, "Archives, Access and AI: Working with Born-Digital and Digitised Archival Collections," Poetry Survival, 2020, https://www.poetrysurvival.com/archives-access-and-ai-working-with-born-digital-and-digitised-archival-collections/.
7. The Royal Society, *Machine Learning: What Do the Public Think?* (The Royal Society, 2017).
8. Kantar Public, *AI PR Survey—Summary of Findings*, 2019.
9. The Royal Society, *Machine Learning*, 2020, https://royalsociety.org/topics-policy/projects/machine-learning/; EU2019FI, "Our Goal Is to Educate 1% of European Citizens in the Basics of AI," Elements of AI, 2020, https://www.elementsofai.com/eu2019fi.
10. Cathy O'Neill, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (London: Penguin, 2016); Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018); Rana Foroohar, *Don't Be Evil: The Case Against Big Tech* (London: Penguin, 2019).

11. Devon Mordell, "Critical Questions for Archives as (Big) Data," *Archivaria* 87 (2019): 140–61.
12. Jason Griffey, "Conclusion," in "Artificial Intelligence and Machine Learning in Libraries, vol. 1," *Library Technology Reports* 55, 2019, 26–28.
13. Sylvester A. Johnson, "Technology Innovation and AI Ethics," in *RLI 299: Ethics of Artificial Intelligence, Research Library Issues* (Association of Research Libraries, 2019), 14; Eileen Jakeway et al., *Machine Learning + Libraries Summit Event Summary* (Library of Congress, 2020), 1.
14. Mary Lee Kennedy, "What Do Artificial Intelligence (AI) and Ethics of AI Mean in the Context of Research Libraries?," in *RLI 299: Ethics of Artificial Intelligence, Research Library Issues* (Association of Research Libraries, 2019), 3–13; William Kilbride, "Nothing About Us Without Us," *DPC Blog* (blog), January 20, 2020, https://dpconline.org/blog/nothing-about-us-without-us.
15. Chris Bourg, "What Happens to Libraries and Librarians When Machines Can Read All the Books?," *Feral Librarian* (blog), March 16, 2017, https://chrisbourg.wordpress.com/2017/03/16/what-happens-to-libraries-and-librarians-when-machines-can-read-all-the-books/; Catherine Nicole Coleman, "Artificial Intelligence and the Library of the Future, Revisited," *Stanford Libraries* (blog), November 3, 2017, https://library.stanford.edu/blogs/digital-library-blog/2017/11/artificial-intelligence-and-library-future-revisited; R. David Lankes, "Decoding AY and Libraries," *R. David Lankes* (blog), July 3, 2019, https://davidlankes.org/decoding-ai-and-libraries/; Annemaree Lloyd, "Chasing Frankenstein's Monster: Information Literacy in the Black Box Society," *Journal of Documentation* 75, no. 6 (2019): 1475–85; Michael Ridley, "Explainable Artificial Intelligence," in *RLI 299: Ethics of Artificial Intelligence, Research Library Issues* (Association of Research Libraries, 2019), 28–46.
16. "Cloud AutoML," Google, Google Cloud, 2020, https://cloud.google.com/automl.
17. Matthias Feurer, Aaron Klein, and Katharina Eggensperger, "Efficient and Robust Automated Machine Learning," *NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems—Volume 2* (December 2015), 2962.
18. Leo Breiman, "Statistical Modeling: The Two Cultures," *Statistical Science* 16, no. 3 (2001): 199–231.
19. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep Learning," *Nature* 521, no. 7553 (2015): 436–44; Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems* (2012), 1097–105; Guenter Muehlberger et al., "Transforming Scholarship in the Archives through Handwritten Text Recognition: Transkribus as a Case Study," *Journal of Documentation* 75, no. 5 (January 1, 2019): 954–76, https://doi.org/10.1108/JD-07-2018-0114.
20. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning Book* (Cambridge, MA: MIT Press, 2016), sec. 6.4.1, https://www.deeplearningbook.org/.
21. David Donobo, "50 Years of Data Science," 2015, 16.
22. Kiri L. Wagstaff, "Machine Learning That Matters," in *Twenty-Ninth International Conference on Machine Learning (ICML)* (Edinburgh, 2012), 529–36.
23. Ali Rahimi and Ben Recht, "Reflections on Random Kitchen Sinks," *Arg Min Blog* (blog), 2017, http://www.argmin.net/2017/12/05/kitchen-sinks/.
24. Paul Voosen, "The AI Detectives," *Science* (2017).
25. Bradley Efron and Trevor Hastie, *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science* (Cambridge, MA: Cambridge University Press, 2016), 447.
26. Efron and Hastie, *Computer Age Statistical Inference*, 447.
27. Jason Brownlee, Machine Learning Mastery, 2020, https://machinelearningmastery.com/; TDS Team, "Machine Learning," Towards Data Science, 2020, https://towardsdatascience.com/machine-learning/home.
28. "Welcome To Colaboratory," Google, Google Colab, 2020, https://colab.research.google.com/notebooks/intro.ipynb.
29. "Notebook Interface," Wikipedia, 2020, https://en.wikipedia.org/wiki/Notebook_interface.
30. Richard Forsyth, "Zoo Data Set," UCI Machine Learning Repository, 2017, https://archive.ics.uci.edu/ml/datasets/Zoo.
31. Ronny Kohavi and Barry Becker, "Census Income Data Set," UCI Machine Learning Repository, 1996, http://archive.ics.uci.edu/ml/datasets/Census+Income.
32. "Amazon Fine Food Reviews," Stanford Network Analysis Project, Kaggle, 2017, https://www.kaggle.com/snap/amazon-fine-food-reviews/.

33. "K-Nearest Neighbors Demo," Stanford Vision and Learning Lab, 2017, http://vision.stanford.edu/teaching/cs231n-demos/knn/; Jinda Shubham, "Decision Trees in Machine Learning," *Becoming Human* (blog), 2018, https://becominghuman.ai/decision-trees-in-machine-learning-f362b296594a.
34. Gaurav Kaushik, "Visualizing Decision Boundaries Builds Intuition About Algorithms," *Medium* (blog), 2018, https://medium.com/cascade-bio-blog/creating-visualizations-to-better-understand-your-data-and-models-part-2-28d5c46e956.
35. "Discovery," The National Archives, 2020, https://discovery.nationalarchives.gov.uk/.
36. Mark Bell, "Machine Learning Club," GitHub, 2020, https://github.com/nationalarchives/MachineLearningClub.
37. "About the Programming Historian," The Programming Historian, 2020, https://programminghistorian.org/en/about; Tim Sherratt, "Welcome to the Wonderful World of GLAM Data!," GLAM Workbench, 2020, https://glam-workbench.github.io/; Nick Ruest et al., "The Archives Unleashed Project: Technology, Process, and Community to Improve Scholarly Access to Web Archives," ArXiv:2001.05399, 2020; Liliana Melgar-Estrada et al., "The CLARIAH Media Suite: A Hybrid Approach to System Design in the Humanities," in *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, CHIIR '19* (Glasgow, Scotland UK: ACM, 2019), 373–77, https://doi.org/10.1145/3295750.3298918.

Bibliography

Alan Turing Institute, The. "Living with Machines." The Alan Turing Institute, 2020. https://www.turing.ac.uk/research/research-projects/living-machines.

Bell, Mark. "Machine Learning Club." GitHub, 2020. https://github.com/nationalarchives/MachineLearningClub.

Bourg, Chris. "What Happens to Libraries and Librarians When Machines Can Read All the Books?" *Feral Librarian* (blog), March 16, 2017. https://chrisbourg.wordpress.com/2017/03/16/what-happens-to-libraries-and-librarians-when-machines-can-read-all-the-books/.

Breiman, Leo. "Statistical Modeling: The Two Cultures." *Statistical Science* 16, no. 3 (2001): 199–231.

Brownlee, Jason. "Machine Learning Mastery." Machine Learning Mastery, 2020. https://machinelearningmastery.com/.

Bunn, Jenny, Yvonne Rogers, Mark Bell, and Jo Pugh. "Workshop on Human-Centred Explainable Artificial Intelligence." *UCL Blog* (blog), 2019. https://blogs.ucl.ac.uk/hexai/.

Coleman, Catherine Nicole. "Artificial Intelligence and the Library of the Future, Revisited." *Stanford Libraries* (blog), November 3, 2017. https://library.stanford.edu/blogs/digital-library-blog/2017/11/artificial-intelligence-and-library-future-revisited.

Donobo, David. "50 Years of Data Science." 2015.

Efron, Bradley, and Trevor Hastie. *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge: Cambridge University Press, 2016.

EU2019FI. "Our Goal Is to Educate 1% of European Citizens in the Basics of AI." Elements of AI, 2020. https://www.elementsofai.com/eu2019fi.

Feurer, Matthias, Aaron Klein, and Katharina Eggensperger. "Efficient and Robust Automated Machine Learning." *NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems—Volume 2* (2015), 2962–70.

Foroohar, Rana. *Don't Be Evil: The Case Against Big Tech*. London: Penguin, 2019.

Forsyth, Richard. "Zoo Data Set." UCI Machine Learning Repository. 2017. https://archive.ics.uci.edu/ml/datasets/Zoo.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning Book*. Cambridge: MIT Press, 2016. https://www.deeplearningbook.org/.

Google. "Cloud AutoML." Google Cloud, 2020. https://cloud.google.com/automl.

———. "Welcome To Colaboratory." Google Colab. 2020. https://colab.research.google.com/notebooks/intro.ipynb.

Goudarouli, Eirini. "Computational Archival Science (CAS): Exploring Data, Investigating Methodologies—Workshop Invitation." *The National Archives Blog* (blog), 2019. https://blog.nationalarchives.gov.uk/computational-archival-science-cas-exploring-data-investigating-methodologies-workshop-invitation/.

Griffey, Jason. “Conclusion.” In “Artificial Intelligence and Machine Learning in Libraries,” 1:26–28. *Library Technology Reports* 55, 2019.
International Council on Archives. “Symposium ‘Archives and AI’ by FAN and the UK National Archives.” ICA, 2019. https://www.ica.org/en/symposium-archives-and-ai-by-fan-and-the-uk-national-archives.
Jaillant, Lise. “Archives, Access and AI: Working with Born-Digital and Digitised Archival Collections.” Poetry Survival, 2020. https://www.poetrysurvival.com/archives-access-and-ai-working-with-born-digital-and-digitised-archival-collections/.
Jakeway, Eileen, Lauren Algee, Laurie Allen, Meghan Ferriter, Jaime Mears, Abigail Potter, and Kate Zwaard. *Machine Learning + Libraries Summit Event Summary*. Library of Congress, 2020.
Johnson, Sylvester A. “Technology Innovation and AI Ethics.” In *RLI 299: Ethics of Artificial Intelligence*, 14–27. Research Library Issues. Association of Research Libraries, 2019.
Kantar Public. *AI PR Survey—Summary of Findings*, 2019.
Kaushik, Gaurav. “Visualizing Decision Boundaries Builds Intuition About Algorithms.” *Medium* (blog), 2018. https://medium.com/cascade-bio-blog/creating-visualizations-to-better-understand-your-data-and-models-part-2-28d5c46e956.
Kennedy, Mary Lee. “What Do Artificial Intelligence (AI) and Ethics of AI Mean in the Context of Research Libraries?” In *RLI 299: Ethics of Artificial Intelligence*, 3–13. Research Library Issues. Association of Research Libraries, 2019.
Kilbride, William. “Nothing About Us Without Us.” *DPC Blog* (blog), January 20, 2020. https://dpconline.org/blog/nothing-about-us-without-us.
Kohavi, Ronny, and Barry Becker. “Census Income Data Set.” UCI Machine Learning Repository. 1996. http://archive.ics.uci.edu/ml/datasets/Census+Income.
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey Hinton. “ImageNet Classification with Deep Convolutional Neural Networks.” *Advances in Neural Information Processing Systems* (2012), 1097–105.
Lankes, R. David. “Decoding AY and Libraries.” *R. David Lankes* (blog), July 3, 2019. https://davidlankes.org/decoding-ai-and-libraries/.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning.” *Nature* 521, no. 7553 (2015): 436–44.
Lloyd, Annemaree. “Chasing Frankenstein’s Monster: Information Literacy in the Black Box Society.” *Journal of Documentation* 75, no. 6 (2019): 1475–85.
Melgar-Estrada, Liliana, Marijn Koolen, Kaspar Beelen, Hugo Huurdeman, Mari Wigham, Carlos Martinez-Ortiz, Jaap Blom, and Roeland Ordelman. “The CLARIAH Media Suite: A Hybrid Approach to System Design in the Humanities.” In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, 373–77. CHIIR ’19. Glasgow, Scotland UK: ACM, 2019. https://doi.org/10.1145/3295750.3298918.
Mordell, Devon. “Critical Questions for Archives as (Big) Data.” *Archivaria* 87 (2019): 140–61.
Muehlberger, Guenter, Louise Seaward, Melissa Terras, Sofia Ares Oliveira, Vicente Bosch, Maximilian Bryan, Sebastian Colutto, et al. “Transforming Scholarship in the Archives through Handwritten Text Recognition: Transkribus as a Case Study.” *Journal of Documentation* 75, no. 5 (January 1, 2019): 954–76. https://doi.org/10.1108/JD-07-2018-0114.
Murphy, Oonagh, and Elena Villaespesa. “The Museums + AI Network.” 2020. https://themuseumsai.network/about/.
National Archives, The. “Digital Strategy 2017–2019.” London: The National Archives, 2017.
———. “Discovery.” The National Archives. 2020. https://discovery.nationalarchives.gov.uk/.
Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.
———. “Annual Digital Lecture: Algorithms of Oppression.” Presented at the Annual Digital Lecture, The National Archives, Kew, 2 April 2019. https://www.youtube.com/watch?v=n3dQUYTN9PA.
O’Neill, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin, 2016.
Programming Historian, The. “About the Programming Historian.” The Programming Historian. 2020. https://programminghistorian.org/en/about.
Rahimi, Ali, and Ben Recht. “Reflections on Random Kitchen Sinks.” *Arg Min Blog* (blog), 2017. http://www.argmin.net/2017/12/05/kitchen-sinks/.
Ridley, Michael. “Explainable Artificial Intelligence.” In *RLI 299: Ethics of Artificial Intelligence*, 28–46. Research Library Issues. Association of Research Libraries, 2019.

Royal Society, The. “Machine Learning.” The Royal Society. 2020. https://royalsociety.org/topics-policy/projects/machine-learning/.
———. ‘Machine Learning: What Do the Public Think?’ The Royal Society. 2017.
Ruest, Nick, Jimmy Lin, Ian Milligan, and Samantha Fritz. “The Archives Unleashed Project: Technology, Process, and Community to Improve Scholarly Access to Web Archives.” ArXiv:2001.05399. 2020.
Sherratt, Tim. “Welcome to the Wonderful World of GLAM Data!” GLAM Workbench, 2020. https://glam-workbench.github.io/.
Shubham, Jinda. “Decision Trees in Machine Learning.” *Becoming Human* (blog), 2018. https://becoming-human.ai/decision-trees-in-machine-learning-f362b296594a.
Stanford Network Analysis Project. “Amazon Fine Food Reviews.” Kaggle, 2017. https://www.kaggle.com/snap/amazon-fine-food-reviews/.
Stanford Vision and Learning Lab. “K-Nearest Neighbors Demo.” 2017. http://vision.stanford.edu/teaching/cs231n-demos/knn/.
TDS Team. “Machine Learning.” Towards Data Science. 2020. https://towardsdatascience.com/machine-learning/home.
University of Oxford. “Visual Geometry Group.” VGG—University of Oxford. 2020. https://www.robots.ox.ac.uk/~vgg/.
Voosen, Paul. “The AI Detectives.” *Science* (2017).
Wagstaff, Kiri L. “Machine Learning That Matters.” In *Twenty-Ninth International Conference on Machine Learning (ICML)*, 529–36. Edinburgh, 2012.
Wikipedia. “Notebook Interface.” Wikipedia. 2020. https://en.wikipedia.org/wiki/Notebook_interface.
Yale University Library. “Digital Humanities Lab.” DH Lab. 2020. https://dhlab.yale.edu/.

Chapter 11

The Role of the Library When Computers Can Read:

Critically Adopting Handwritten Text Recognition (HTR) Technologies to Support Research

Melissa Terras

Introduction

Computational approaches to processing and searching images of historical manuscripts by handwritten text recognition (HTR) is one of the most promising machine learning approaches for academic research in the humanities, having the potential to transform access to our written past for the use of researchers, institutions, and the general public. This chapter surveys the current use of HTR in the library sector, highlighting major tools currently in use and activities being undertaken by academic and research libraries. Using Transkribus as a case study, the chapter provides examples of where libraries have successfully deployed HTR and focuses on emerging issues for incorporating the application and results of HTR into a digitisation workflow, including documentation, results delivery, and sustainability. The chapter considers how HTR can be best deployed to support researchers, including the need for transparency, training, and data infrastructure. Although HTR technology is now reasonably mature, academic libraries need to adopt this machine learning (ML) technique in a critical way, signposting the data in a way that explains its creation and allows its embedding into historical practice to best support their user communities.

Computationally Reading Documents: OCR, HTR, and Mass Digitisation

Libraries and archives have invested heavily in mass digitisation of their print and manuscript collections over the past thirty years.[1] The textual content of the resulting digital images, however, was only recently available to those with the resources to manually transcribe individual passages.[2] Manual transcription is an approach that does not scale across larger collections of documents given the costs associated with employing researchers or the setup and monitoring costs associated with working with volunteers (although "crowdsourcing" such volunteer labour using online mechanisms can be cost-effective in the long term).[3] It has long been an aim of computer scientists, librarians, archivists, and curators to be able to generate accurate machine-readable transcriptions of their holdings, with the accepted understanding that doing so will enable the "key functional elements of large databases of print—easily readable texts and full-text searching."[4] Including handwritten material "promises to yet again extend and revolutionize the study of historical handwritten documents,"[5] allowing both searching at a scale impossible to the offline human reader and the use of advanced mining, analysis, and visualisation techniques.[6] This brings with it the possibility of providing new untold insights to benefit the research community, utilising (and democratising access to) the vast volume of digital images of manuscripts now produced by the heritage sector.

Both optical character recognition (OCR, the conversion of images of typed or printed text into machine-encoded format) and handwritten text recognition (HTR, the use of computational technologies to interpret text from handwritten sources and to produce it in machine-encoded format) have long histories, stretching back as early as the nineteenth century.[7] OCR is now a standard approach, routinely embedded within digitisation workflows and digital library programs,[8] with the resulting datasets allowing for searching across massive repositories of digitised text. There are various OCR solutions available (which themselves use AI), including those which are commercial (ABBYY FineReader, Kofax OmniPage Standard, Adobe Acrobat DC)[9] and open source (Tesseract, now owned by Google).[10] Issues with accuracy, however, remain, particularly with formats that deviate from clearly printed text, such as the inky spread of newsprint or more complex fonts and page presentation, and that can affect resulting further analysis of the material.[11] Recent developments in HTR, involving both machine and deep-learning approaches, mean it is now possible to generate machine-processable text directly from digitised images of handwritten (or complex print) material.[12] Improved accuracy rates have increased discoverability and the potential to undertake new and novel research at scale.[13] As a result, HTR is one of the few applications of artificial intelligence in the cultural and heritage sector that has reached relative maturity and is now being applied by digitisation units and scholarly projects across the academic library sector. Use is far from standardised and there has been next to no research on how best practices can be undertaken in storing, sharing, and explaining HTR-generated content in this field.

The Current HTR Landscape for Libraries

Libraries are at a point where they must choose if they want to trial this potentially useful technology within their heritage digitisation workflows. They can work in tandem with computer scientists and generate their own bespoke HTR approaches for the material in question (there being a long history of this research-led approach to reading historical texts),[14] or they can reuse the outputs of these projects. For example, the In Codice Ratio project is developing "tools to support content analysis and knowledge discovery from large collections of historical documents," concentrating on the collections of the Vatican Secret Archives.[15] Using a convolutional neural network classifier[*] and statistical language models to generate the most likely transcript, it also engages palaeographers in crowd-sourcing training data. Likewise, the Monk system has been developed by the University of Groningen using AI methods for accessing historical archive collections that are difficult to process by traditional OCR methods—for example, due to their historical character types or due to the fact that the material is handwritten. The system consists of two major components: (1) a setup for the storage and web-based annotation of scanned page images and parts thereof, and (2) a set of (handwriting and text) recognition algorithms as well as retrieval and search methods.[16]

These systems work with a variety of human languages and across temporalities: the PYTHIA project is the first ancient text restoration model that recovers missing characters from damaged Greek inscriptions using deep neural networks;[17] and The Center for Open Data in the Humanities in Tokyo is developing approaches and datasets for training and reading different types of Japanese script, such as deep-learning techniques

* A convolutional neural network, or CNN, is a particular type of deep neural network (an artificial neural network with multiple layers between input and output) that is mostly deployed in the analysis of digital images. Inspired by the design of the animal visual cortex, they provide an efficient pattern recognition approach. See Saad Albawi, Mohammed Tareq Abed, and Saad Al-Zawi, *Understanding of a convolutional neural network*, 2017.

to be used in conjunction with the Kuzushiji Dataset of Pre-Modern Japanese Text.[18] The projects mentioned here are only an indicative few, which are part of a large and expanding community of computational researchers developing their own HTR solutions as well as publishing their code and results for others to evaluate and reuse with emerging benchmarks and best-practice guidelines.[19] Establishing such partnerships, however, requires resources, expertise, and confidence in the sustainability of chosen systems. Libraries should have a robust plan as to how they will manage and maintain both the code and the resulting datasets if these in-house solutions are adopted and applied to mass-digitised content.

Libraries can turn to the solutions being offered by publishers and established technology companies, although the processes used to generate HTR can be opaque. For example, Adam Matthew Digital "is the first publisher to utilise artificial intelligence to offer handwritten text recognition (HTR) for its handwritten manuscript collections" at time of writing, offering full-text search of transcriptions of seven major archival collections that have been processed with their in-house HTR.[20] Institutions can now license access to their Quartex document management system, which is the "only platform with built-in HTR, making manuscripts searchable."[21] Likewise, Gale is offering full-text searching for two of its collections processed with HTR, although not providing access to a platform for others to upload their content.[22] Google Arts and Culture recently launched Fabricius, using AI to help translate hieroglyphs, and they are also continuing to expand API for developers wishing to detect handwriting in images.[23] This potential relationship with corporate publishers and technological giants is therefore one to closely watch as they continue to develop tools to apply HTR to mass-digitised content. As with relationships with all digital commercial entities, care should be given to copyright; image licensing; mechanisms for storing, sharing, and long-term archiving of both input and output data; and access to explanations of the algorithms involved to better understand how the data is processed.

Transkribus as an HTR Solution for Libraries

Between 2011 and 2019, a large consortium of EU-funded researchers led by the University of Innsbruck developed a machine-learning approach for automatic generation of transcripts from digitised images of historical handwritten text, which resulted in software—Transkribus[*]—which is now capable of generating transcriptions with up to 98 percent accuracy.[24] At the time of writing, Transkribus has forty-one thousand registered users, including individuals and major libraries, archives, and museums worldwide. Over one thousand users actively use the software each week, with more than a million images uploaded every month for processing and over six thousand HTR models now generated in total by the community. From July 1, 2019, the platform has been operated and further

* See: https://transkribus.eu/Transkribus/.

developed by the European cooperative READ-COOP, a mechanism planned to sustain and grow the Transkribus infrastructure beyond the end of its grant-financed period built around a cooperative economic model.[25]

Transkribus uses deep neural network machine learning technology. Once images of manuscripts are uploaded, layout analysis tools segment them into lines (the Transkribus graphical user interface (GUI) contains both automatic and manual segmentation tools, allowing user correction), before each line is transcribed. A training process—on approximately fifteen thousand transcribed words or seventy-five pages of handwritten script—generates what is known as an HTR *model* for recognising text written in one hand. Users can either apply a model created (and sometimes openly published) by another project or train up their own, which is then used to generate transcriptions.[26] In the best cases, HTR can produce automated transcriptions of handwritten material with a character error rate (CER) of below 5 percent, meaning 95 percent of characters are correct, and if used on printed material, that CER can reach 1 to 2 percent. Users can then work with the Transkribus GUI to correct and improve the generated transcripts to compile new, improved, models; this creates a feedback loop that improves the efficacy of the underlying neural network, benefiting future users. Once HTR has been completed, the user is free to take the resulting transcriptions and use them in any way they feel appropriate, such as inclusion in digital scholarly editions, used as source data for further linguistic or semantic analysis, or ingested into digital library content management systems as a finding aid used in conjunction with keyword searching.[27]

An early user of Transkribus was the Bentham Project at University College London, which trained models on the writing of the philosopher and jurist Jeremy Bentham (1748–1832) in order to transcribe his vast personal archive. The initial models were trained on crowdsourced transcriptions of Bentham's manuscript material.[28] With forty-one thousand users in at least fifty-three countries worldwide, there are a broad variety of institutions that have adopted the platform for a range of complex projects.[29] For example, transcribing a large set of Michel Foucault's reading notes, including citations, references, and comments; increasing the accessibility and usability of the archive of the cloister of the Poor Clares St.-Elisabethsdal in Boxtel (1390–1719); and retro-digitizing and automatically structuring the large bibliography collection of the *Internationale Bibliographie der Lexikographie* by Herbert Ernst Wiegand.[30] These examples indicate the range of potential uses on digitised academic library collections while also presenting a resource-saving opportunity and removing barriers caused by not having a budget to employ staff to transcribe these archival manuscript materials themselves.

The Transkribus platform is now at a crucial moment: having built a working platform, the grant-funded period has now ended, and to maintain the infrastructure, an income stream must be created. The platform will be transferred to a paid-for model in late 2020 and, therefore, library projects will have to include a budget for its use in their plans. There are other criticisms of Transkribus, the main one vocalized being the fact that not all of its processes, algorithms, and models are published, and so it is not fully subscribing to the Open Science principles.[31] Via the READ-COOP, Transkribus continues to work with its growing user base in order to provide continued access to this tool for the wider cultural heritage community. There are regular meetings and lively Facebook user forums (both official and unofficial) to assist when applying HTR via Transkribus to digitised content. It has rapidly become the most recommended tool for HTR (and AI, by extension) within the cultural heritage sector.

Critically Embedding HTR Into Libraries

Although there are considerable savings in time and resources in using HTR to generate transcriptions of historical manuscripts, HTR is not a panacea. If it is to be successfully used to increase access to information within and usability of handwritten textual material, it needs to be embedded into both digitisation workflows within libraries and other institutions as well as public-facing digital library infrastructures. However, there is little consideration to date of how HTR can be built into service-level provision of digitisation within academic libraries that adopt it or the type of messaging and communication that will be necessary to allow users to embed HTR-generated data into their research practice.

HTR's use is dependent on the availability of digitised content, and that itself has long been known to be a costly and complex endeavour, which often has ethical and legal implications, including the necessary navigation of copyright and related permissions.[32] HTR has the potential to extend the diversity of materials available to researchers, but only if libraries engage with them at the digitisation phase, ensuring a plurality of voices and that archives are entering the digitisation pipeline, which has not been the case to date.[33]

While there are efficiencies in employing HTR to generate transcripts of images of handwritten text, rather than employing researchers to do so, the training of models and generating of transcripts by HTR is not fully automated. For optimal results, there needs to be full engagement with the feedback loops built into the process; therefore, it does still take resources to deploy HTR successfully. As well as finding resources, libraries need to plan ahead regarding how HTR can become embedded into existing digitisation and information workflows. It is essential that projects establish a data-management plan for the resulting HTR-generated transcripts. Integrating HTR into digitisation processes (managed by platforms such as Goobi)[34] and embedding the results of HTR within content management systems (CMS) are essential steps to reap the benefits of this technology and promote the discoverability of the results, thus ensuring that the data is preserved and following normal institutional digital-curation practices. While there are various approaches to handling OCR-generated data that can be adopted, there are not yet any *de facto* standards for even the basics in HTR data handling, including industry file naming standards for this content, never mind more sophisticated workflows, such as persistent identifier generation, links to other infrastructural frameworks such as IIIF,[*] or standards

* The International Interoperability Framework (IIIF) aims to define a set of common standards and application programming interfaces that support interoperability between image repositories, promote best practice in this area, and become more adopted in the GLAM sector; see https://iiif.io.

for openly publishing datasets (although projects such as EScripta are starting to explore these opportunities, and should be watched).[35]

In addition, most user-facing CMSs are not yet configured to allow this additional computer-generated textual information to sit alongside high-resolution images and to enable discoverability by full-text searching. It is also unclear where, in now-standard library and archival metadata structures, these AI-generated transcripts should be stored or if they are covered by non-print legal deposit regulations for mandated ingestion into national digital library repositories. It is not known how best to manage the reporting of any errors, in order to correct and improve HTR-generated content, in a way that is both transparent and scalable (although there are parallels with mass OCR correction).[36] Libraries, therefore, need to carefully consider both the ingestion, processing, and output of HTR processes to ensure that it is employed in a useful and sustainable manner alongside existing infrastructure.

In addition, there are issues regarding information literacy. There is a need to highlight to users of library and archival systems which information fields have been generated (and checked) by trained professionals in the sector, which have been automatically created by algorithms, and which may be crowdsourced from other users as part of the feedback process. Researchers need to be made aware of how datasets are created from digitisation to dataset, and there has been little work on the transparency of HTR tools and techniques, a criticism that can be levelled at the major technological providers (including Transkribus). Academic libraries must address these issues in a clearly explained manner in order to support their research users adequately. There has been next to no research on the users of HTR-generated historical content and the implications this technology has for scholarly work. If we are to see HTR-generated datasets used as new source material to underpin novel research, we must be able to explain their provenance so that the resulting datasets can be trusted as a scholarly source. Libraries will have to provide training and support for researchers to understand appropriate methodologies to get the best results from the emerging, relatively large-scale datasets and to provide new infrastructures to host and support these new "collections as data."[37] The claims made that HTR has the potential to revolutionise scholarship should be viewed with interest by libraries. They will have a supporting function in the creation, hosting, and delivery of such large-scale transcripts and in supporting the user community to make the most of the opportunities therein.

Beyond scholarly applications, there are wider environmental issues that need to be addressed with the use of high-performance computing. It can take several days for a server farm to train an HTR model, which has implications for energy usage and carbon emissions. Many institutions are starting to think holistically about their use of such resources and mitigate against them with schemes such as carbon-offsetting, which will need to be considered and factored into project plans.[38]

As a result, HTR may be a maturing technology from an algorithmic perspective, but procedurally, libraries have to adopt it critically to help these issues settle into best practices for the sector. Given the newness of the approach and the tools involved, we are in a parallel position to where libraries were in the 1990s when digitisation at scale first became affordable, possible, and practical, and where the reports of individual libraries on their attempts to operationalise this approach were essential to establishing sector best practices. At this stage in the HTR journey, it behooves projects and institutions adopting this technology to consciously report on the decisions, practices, and protocols that are necessary within the digitisation and usability pipeline. This will allow others to learn from previous implementations to help the sector work toward a user-centric best practice of this promising AI technology and to aid research into the human notes of the past from a conscious, informed position.

Conclusion

Academic libraries have an important role to play now that computers can read handwritten text in selecting diverse material to be digitised and supporting a plurality of voices in our digital cultural heritage landscape. Libraries must decide where to apply HTR to generate searchable and processable outputs from these sources; establish best practices to ensure that these data sources are sustainable, findable, and useable; and support the research community in accessing, analysing, and reporting on their content. In order to make the most of the opportunities recently presented by HTR technology, those in the cultural heritage sector using it (however experimentally) should report, communicate, and discuss both methods and results with others similarly interested to contribute to a convergence of approaches which will, eventually, become standardised best practice. It is essential that this new machine-learning technology be harnessed by those wishing to increase access to content held within manuscript material that will be of interest to researchers. Doing so means a level of understanding, engagement, and control to establish where HTR can contribute to the work of academic researchers in their understanding of the past and to make processes transparent and understandable. Libraries are expertly placed to be the nexus that can support and encourage the novel research questions, approaches, and, ultimately, outputs that give handwritten text recognition its transformational potential. As this approach continues to develop, libraries can also frame and explain how the use of this technology may change and expand existing historical scholarly practice given certain ventures into these new AI-generated vistas.

Conflict of interest statement: Melissa Terras is on the Board of Directors of the READ COOP. She does not and will not benefit financially from this, either now or in the future, due to the cooperative financial structures in place.

Endnotes

1. Melissa Terras, "The Rise of Digitization: An Overview," in *Digitisation Perspectives*, ed. Ruth Rikowski (Leiden, Netherlands: Brill, 2011), 1–20; Natasha Stroeker and René Vogels, "Survey report on digitisation in European cultural heritage institutions 2012," Panteia, Netherlands, on behalf of the ENUMERATE Thematic Network, 2014.
2. John Edwards, "Easily adaptable handwriting recognition in historical manuscripts" (PhD Dissertation, University of California, Berkeley, CA, 2007).
3. Tim Causer, Kris Grint, Anna-Maria Sichani, and Melissa Terras, "'Making such bargain': Transcribe Bentham and the quality and cost-effectiveness of crowdsourced transcription," *Digital Scholarship in the Humanities* 33, no. 3 (2018): 467–87.

4. Laura Estill and Michelle Levy, "Chapter 12: Evaluating Digital Remediations of Women's Manuscripts," Digital Studies/Le champ numérique, Vol. 6, 2016, http://doi.org/10.16995/dscn.12.
5. Estill and Levy, "Chapter 12: Evaluating Digital Remediations."
6. Kristen Schuster and Stuart Dunn, *Routledge International Handbook of Research Methods in Digital Humanities* (Oxfordshire, UK: Routledge, 2020).
7. Herbert Schantz, *The History of OCR, Optical Character Recognition*, Manchester Center, VT: Recognition Technologies Users Association, 1982.
8. Karmen Sotošek, "Best practice examples in library digitisation," Europeana Travel, D2.2, 011, https://pro-beta.europeana.eu/files/Europeana_Professional/Projects/Project_list/EuropeanaTravel/Deliverables/D2.2%20Best%20practice%20examples%20in%20library%20digitisation.pdf.
9. ABBYY FineReader, 2020, https://pdf.abbyy.com/; Kofax Omnipage Standard, 2020, https://www.kofax.com/Products/omnipage/standard; Adobe Acrobat DC, Adobe, 2020, https://acrobat.adobe.com/uk/en/acrobat/pricing.html.
10. Tesseract-OCR, 2020, https://github.com/tesseract-ocr.
11. Simon Tanner, Trevor Muñoz, and Pich Hemy Ros, "Measuring mass text digitization quality and usefulness: Lessons Learned from Assessing the OCR Accuracy of the British Library's 19th Century Online Newspaper Archive," *D-lib Magazine* 15, no. 7/8 (2009), http://www.dlib.org/dlib/july09/munoz/07munoz.html; Ryan Cordell, "'Q i-jtb the Raven': Taking Dirty OCR Seriously," *Book History* 20, No. 1 (2017): 188–225.
12. Byron Bezerra et al., *Handwriting: Recognition, Development and Analysis* (New York: Nova Science Publishers, 2017).
13. Estill and Levy, "Chapter 12: Evaluating Digital Remediations."
14. Melissa M. Terras, *Image to Interpretation: An Intelligent System to Aid Historians in Reading the Vindolanda Texts* (Oxford University Press, 2006).
15. Donatella Firmani et al., "Towards Knowledge Discovery from the Vatican Secret Archives. In Codice Ratio-Episode 1: Machine Transcription of the Manuscripts," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, 263–72.
16. Monk, *Lambert Schomaker, Lifelong learning for text retrieval and recognition in historical handwritten document collections*, 2019, https://www.ai.rug.nl/~lambert/Monk-collections-english.html.
17. Assael Yannis, Thea Sommerschield, and Jonathan Prag, "Restoring ancient text using deep learning: a case study on Greek epigraphy," arXiv preprint arXiv:1910.06262, 2019.
18. Tarin Clanuwat, Alex Lamb, and Asanobu Kitamoto, "Kuronet: Pre-modern Japanese kuzushiji character recognition with deep learning," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 2019, 607–14.
19. Joan Andreu Sánchez, Verónica Romero, Alejandro H. Toselli, Mauricio Villegas, and Enrique Vidal, "A set of benchmarks for handwritten text recognition on historical documents," *Pattern Recognition* 94 (2019): 122–34.
20. "Artificial intelligence transforms discoverability of handwritten manuscripts," Adam Matthew Digital, 2020, https://www.amdigital.co.uk/products/handwritten-text-recognition.
21. "Features," Quartex, 2020, https://www.quartexcollections.com.
22. "An essential primary source archive for researching the history of Hong Kong in the context of Modern China and the British Empire in Asia," Gale, 2020, https://www.gale.com/intl/c/china-and-the-modern-world-hongkong-britain-china.
23. "What is Fabricius?," Google Arts and Culture, 2020; "Detect handwriting in images," Google Cloud, 2020.
24. Guenter Muehlberger et al., "Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study," *Journal of Documentation* (2019), https://doi.org/10.1108/JD-07-2018-0114.
25. READ-COOP, "Revolutionizing Access to Handwritten Documents," 2020, https://readcoop.eu.
26. Philip Kahle, Sebastian Colutto, Günter Hackl, and Günter Mühlberger, "Transkribus-A service platform for transcription, recognition and retrieval of historical documents," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 4 (2017), 19–24.
27. Muehlberger et al., "Transforming scholarship."

28. Tim Causer and Melissa Terras, "'Many hands make light work. Many hands together make merry work': Transcribe Bentham and Crowdsourcing manuscript collections," in *Crowdsourcing Our Cultural Heritage*, ed. Mia Ridge (Farnham, UK: Ashgate Publishing, 2014), 57–88.
29. Melissa Terras, Tweet, 2019.
30. Vincent Ventresque, Arianna Sforzini, and Marie-Laure Massot, "Transcribing Foucault's handwriting with Transkribus," *Journal of Data Mining & Digital Humanities* (2019), https://jdmdh.episciences.org/5218/pdf; Geertrui Van Synghel, "'Sine clausura': Unlocking the archive of the cloister of the Poor Clares St.-Elisabethsdal in Boxtel (1390-1719)," *Franciscan Studies* 77, no. 1 (2019): 89–110; David Lindemann, Mohamed Khemakhem, and Laurent Romary, "Retro-digitizing and Automatically Structuring a Large Bibliography Collection," European Association for Digital Humanities (EADH) Conference, EADH, December 2018, Galway, Ireland, https://hal.archives-ouvertes.fr/hal-01941534/document.
31. "What is Open Science? Introduction," Foster, 2020, https://www.fosteropenscience.eu/node/1420.
32. Lorna Hughes, *Digitizing Collections: Strategic Issues for the Information Manager* (London: Facet Publishing, 2004).
33. Nanna Thylstrup, *The Politics of Mass Digitization* (Cambridge: MIT Press, 2019).
34. "Goobi for the workflow," Intranda, 2020, https://www.intranda.com/en/digiverso/goobi/workflow/.
35. Peter Stokes, Benjamin Kiessling, Robin, Tissot, Daniel Stökl, and Ben Ezra, "eScripta: A New Digital Platform for the Study of Historical Texts and Writing," Digital Humanities 2019, Utrecht University, July 2019, https://dev.clariah.nl/files/dh2019/boa/0322.html.
36. Rose Holley, "Many hands make light work: Public collaborative OCR text correction in Australian historic newspapers," National Library of Australia, 2009, https://www.nla.gov.au/content/many-hands-make-light-work-public-collaborative-ocr-text-correction-in-australian-historic.
37. Thomas Padilla, Laurie Allen, and Hannah Frost, "Always Already Computational: Collections as Data," Zenodo, 2019, http://doi.org/10.5281/zenodo.3152935.
38. Loïc Lannelongue, Jason Grealey, and Michael Inouye, "Green Algorithms: Quantifying the carbon emissions of computation," arXiv preprint arXiv:2007.07610, 2020.

Bibliography

ABBYY FineReader PDF. 2020. https://pdf.abbyy.com/.

Adam Matthew Digital. "Artificial intelligence transforms discoverability of handwritten manuscripts." 2020. https://www.amdigital.co.uk/products/handwritten-text-recognition.

Assael, Yannis, Thea Sommerschield, and Jonathan Prag. "Restoring ancient text using deep learning: a case study on Greek epigraphy." arXiv preprint arXiv:1910.06262. 2019.

Adobe. "Adobe Acrobat DC Plans and Pricing." 2020. https://acrobat.adobe.com/uk/en/acrobat/pricing.html.

Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a convolutional neural network." In *2017 International Conference on Engineering and Technology (ICET)* (2017): 1–6.

Bezerra, Byron, et al., *Handwriting: Recognition, Development and Analysis*. New York: Nova Science Publishers, 2017.

Causer, Tim, and Melissa Terras. "'Many hands make light work. Many hands together make merry work': Transcribe Bentham and Crowdsourcing manuscript collections." In *Crowdsourcing Our Cultural Heritage*, edited by Mia Ridge. Farnham, UK: Ashgate, 2014, 57–88.

Causer, Tim, Kris Grint, Anna-Maria Sichani, and Melissa Terras. "'Making such bargain': Transcribe Bentham and the quality and cost-effectiveness of crowdsourced transcription." *Digital Scholarship in the Humanities* 33, no. 3 (2018): 467–87.

Clanuwat, Tarin, Alex Lamb, and Asanobu Kitamoto. "Kuronet: Pre-modern Japanese kuzushiji character recognition with deep learning." In *2019 International Conference on Document Analysis and Recognition (ICDAR)*. IEEE (2019), 607–14.

Cordell, Ryan. "'Q i-jtb the Raven': Taking Dirty OCR Seriously." *Book History* 20, No. 1 (2017): 188–225.

Edwards, John Alexander. "Easily adaptable handwriting recognition in historical manuscripts." PhD Dissertation, University of California, Berkeley, CA, 2007. http://www2.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-76.pdf.

Estill, Laura, and Michelle Levy. “Chapter 12: Evaluating digital remediations of women’s manuscripts.” Digital Studies/Le champ numérique, Vol. 6. 2016. http://doi.org/10.16995/dscn.12.
Firmani, Donatella, Marco Maiorino, Paolo Merialdo, and Elena Nieddu. “Towards Knowledge Discovery from the Vatican Secret Archives. In Codice Ratio-Episode 1: Machine Transcription of the Manuscripts.” In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2018), 263–72.
Foster, “What is Open Science? Introduction.” 2020. https://www.fosteropenscience.eu/node/1420.
Gale. “An essential primary source archive for researching the history of Hong Kong in the context of Modern China and the British Empire in Asia.” 2020. https://www.gale.com/intl/c/china-and-the-modern-world-hongkong-britain-china.
Google Arts and Culture. “What is Fabricius?” 2020. https://artsexperiments.withgoogle.com/fabricius/en.
Google Cloud. “Detect handwriting in images.” AI and Machine Learning Products. 2020. https://cloud.google.com/vision/docs/handwriting.
Holley, Rose. “Many hands make light work: Public collaborative OCR text correction in Australian historic newspapers.” National Library of Australia. 2009. https://www.nla.gov.au/content/many-hands-make-light-work-public-collaborative-ocr-text-correction-in-australian-historic.
Hughes, Lorna M. *Digitizing Collections: Strategic Issues for the Information Manager*. London: Facet Publishing, 2004.
Intranda. “Goobi for the workflow.” 2020. https://www.intranda.com/en/digiverso/goobi/workflow/.
Kahle, Philip, Sebastian Colutto, Günter Hackl, and Günter Mühlberger. “Transkribus—A service platform for transcription, recognition and retrieval of historical documents.” In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 4, IEEE (2017), 19–24.
Kofax Ominpage Standard. 2020. https://www.kofax.com/Products/omnipage/standard.
Lindemann, David, Mohamed Khemakhem, and Laurent Romary. “Retro-digitizing and Automatically Structuring a Large Bibliography Collection.” European Association for Digital Humanities (EADH) Conference, EADH, December 2018, Galway, Ireland. https://hal.archives-ouvertes.fr/hal-01941534/document.
Lannelongue, Loïc, Jason Grealey, and Michael Inouye. “Green Algorithms: Quantifying the carbon emissions of computation.” arXiv preprint arXiv:2007.07610, 2020.
Monk. “Indexed books and overview.” https://www.ai.rug.nl/~lambert/Monk-collections-english.html.
Muehlberger, Guenter, Louise Seaward, Melissa Terras, et al. “Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study.” *Journal of Documentation* (2019). https://doi.org/10.1108/JD-07-2018-0114.
Padilla, Thomas, Laurie Allen, Hannah Frost, et al. “Always Already Computational: Collections as Data.” Zenodo. 2019. http://doi.org/10.5281/zenodo.3152935.
Quartex, “Features.” 2020. https://www.quartexcollections.com.
READ-COOP. “Revolutionizing Access to Handwritten Documents.” 2020. https://readcoop.eu.
Sánchez, Joan Andreu, Verónica Romero, Alejandro H. Toselli, Mauricio Villegas, and Enrique Vidal. “A set of benchmarks for handwritten text recognition on historical documents.” *Pattern Recognition* 94 (2019): 122–34.
Schantz, Herbert H. “The history of OCR, optical character recognition.” Manchester Center, VT: Recognition Technologies Users Association. 1982.
Schomaker, Lambert. “Lifelong learning for text retrieval and recognition in historical handwritten document collections.” Forthcoming in *Handwritten Historical Document Analysis, Recognition, and Retrieval, State of the Art and Future Trends*. arXiv preprint arXiv:1912.05156. 2019.
Schuster, Kristen, and Stuart Dunn, eds. *Routledge International Handbook of Research Methods in Digital Humanities*. Oxfordshire, UK: Routledge, 2020.
Stroeker, Natasha, and René Vogels. “Survey report on digitisation in European cultural heritage institutions 2012.” Panteia, Netherlands. On behalf of the ENUMERATE Thematic Network, 2014.
Sotošek, Karmen Štular. “Best practice examples in library digitisation.” Europeana Travel, D2.2. 2011. https://pro-beta.europeana.eu/files/Europeana_Professional/Projects/Project_list/EuropeanaTravel/Deliverables/D2.2%20Best%20practice%20examples%20in%20library%20digitisation.pdf.
Stokes, Peter, Benjamin Kiessling, Robin, Tissot, Daniel Stökl, and Ben Ezra. “eScripta: A New Digital Platform for the Study of Historical Texts and Writing.” Digital Humanities 2019, Utrecht University. July 2019. https://dev.clariah.nl/files/dh2019/boa/0322.html.

Tanner, Simon, Trevor Muñoz, and Pich Hemy Ros. "Measuring mass text digitization quality and usefulness: Lessons Learned from Assessing the OCR Accuracy of the British Library's 19th Century Online Newspaper Archive." *D-lib Magazine* 15, no. 7/8 (2009). http://www.dlib.org/dlib/july09/munoz/07munoz.html.

Terras, Melissa M. "The Rise of Digitization: An Overview," in *Digitisation Perspectives*, ed. Ruth Rikowski. Leiden, Netherlands: Brill, 2011, 1–20.

———. *Image to Interpretation: An Intelligent System to Aid Historians in Reading the Vindolanda Texts*. Oxford University Press, 2006.

———. Tweet: "So @Transkribus now has 22k users. 3k of those logged in over the past 3 months, in 31k sessions. Altho the location of 37% of those are indeterminable (gmail, hotmail, etc), I plotted the rest: according to their domains, we have active users in 53 countries worldwide!" July 10, 2019. https://twitter.com/melissaterras/status/1148859050735624192/photo/3.

Tesseract-OCR. 2020. https://github.com/tesseract-ocr.

Thylstrup, Nanna Bonde. *The Politics of Mass Digitization*. Cambridge: MIT Press, 2019.

Transkribus. 2020. https://transkribus.eu/Transkribus/.

Van Synghel, Geertrui. "'Sine clausura': Unlocking the archive of the cloister of the Poor Clares St.-Elisabethsdal in Boxtel (1390–1719)." *Franciscan Studies* 77, no. 1 (2019), 89–110.

Ventresque, Vincent, Arianna Sforzini, and Marie-Laure Massot. "Transcribing Foucault's handwriting with Transkribus." *Journal of Data Mining & Digital Humanities* (2019). https://jdmdh.episciences.org/5218/pdf.

Chapter 12

Using IBM Watson for Discovery and Research Support:

A Library–Industry Partnership at Auburn University

Aaron Trehub and Ali Krzton

Introduction

Researchers at the Auburn University Libraries are collaborating with cross-campus units and private companies to explore the application of artificial intelligence and machine learning (AI/ML) tools to scholarly repositories, structured and unstructured datasets, and the open web. The goal is twofold: to develop hands-on AI/ML expertise in the libraries and to better position the libraries to support and participate in Auburn University's research activities.

This chapter describes an exploratory project the libraries are working on with the IBM Research Triangle Park Center for Advanced Studies (IBM RTP-CAS) on using the IBM Watson Studio of AI/ML services to build natural-language query interfaces for scholarly repositories and datasets in specific subject domains. Since 2017, the libraries have been providing high-level information technology support and subject-matter expertise to the Military REACH Project, which is based in the College of Human Sciences (CHS) at Auburn University and funded by the US Department of Agriculture and the

US Department of Defense. The mission of the Military REACH Project is to support US military families and family readiness by translating academic research and other resources into practical applications. A key component of the project is the design and development of a publicly accessible, easy-to-navigate library of research publications and other resources on military life and family health, delivered through a user-friendly website. With IBM's help, the Watson suite of tools and services is being used to build an AI/ML-powered query interface and recommendation system for the Military REACH Library and public datasets in the same field.

In this chapter, the authors discuss the larger library and professional context for this project, its background and rationale, the deliverables they are producing, and the technical, logistical, and administrative challenges that they have encountered in this effort. The goal is to provide other academic libraries with a model for embarking on similar projects and a clear understanding of the benefits and challenges involved.

Setting the Context: AI/ML and Academic Libraries

The past five years have seen an increasing number of articles, technical reports, white papers, and webinars on the implications of AI/ML for libraries in general and academic libraries in particular. Early pieces on AI/ML and libraries were mostly speculative or theoretical in nature. They focused on what AI/ML might mean for academic libraries and information work in general. More recently, the discussion has expanded to include working examples of how AI/ML is being used in library applications and the establishment of dedicated AI laboratories in academic libraries.

Writing in 2016, Peter Fernandez of the University of Tennessee Libraries identified seven library functions that will be affected or enabled by AI: discovery, cataloging and metadata creation, translational reference, interpretation collection analysis and development, and storage and inventory management.[1] UK-based researchers Andrew Cox, Stephen Penfield, and Sophie Rutter covered some of the same territory in a 2019 article on "the intelligent library."[2] They interviewed thirty-three stakeholders inside and outside the academic library community on the implications of AI for academic libraries, issues arising from the use of AI applications, and the role of academic libraries in supporting and using AI. Among the possible library roles identified by interviewees were: procuring or creating content for AI/ML services; procuring or designing AI tools; performing data curation, quality control, and analysis; designing data infrastructure; teaching critical data literacy; serving as navigators to the new information environment; and writing AI algorithms.[3]

Kenning Arlitsch and Bruce Newell, however, argued in a 2017 article that AI/ML will almost certainly have substantial disruptive effects on library operations and library employment. Citing a much-referenced paper on computerization and the future of employment by Frey and Osborne, Arlitsch and Newell observed that certain library tasks or operations were at risk of being replaced by robots.[4] They argued that librarians need to start acquiring "quantitative and analytical skills to learn the value of big data" and "make the machines work for us."[5] They left open the question of whether this is a realistic goal for the current generation of librarians, whose training and career paths have largely been determined by traditional MLIS programs.

Perceptions of and attitudes toward AI/ML in the library community have shifted from complacency to urgency in the past several years. In a 2018 article on the results of a survey they conducted, Barbara Wood and David Evans at Kennesaw State University identified "an overwhelming sense of complacency among librarians" about the disruptive effects of AI. Although 56 percent of the more than 300 librarians who responded to the survey answered in the affirmative when asked, "Do you think supercomputers like [IBM] Watson will have a transformative effect on librarianship?", almost 44 percent believed that AI would have no or not much effect on the profession. At the same time, most librarians surveyed put the probability of AI/ML solutions like IBM Watson being used in libraries at 50 percent by 2027 and 90 percent by 2047.[6]

The combination of big data and AI/ML raises serious ethical and policy questions in almost every area of society. Librarianship and information management are deeply implicated in these questions. Library observers have identified a number of ethical issues and dangers associated with the irresponsible or maladroit use of AI/ML for library applications. The main ones have to do with promoting or contributing to misinformation, "fake news," algorithmic bias, violations of patron privacy and the possibility of high-tech surveillance, and "perpetuat[ing] existing forms of structural inequality."[7] In 2019, the Association of Research Libraries devoted a special report to libraries and the ethics of artificial intelligence.[8]

A growing number of libraries are moving beyond AI/ML theory to practice. Examples include the Stanford University Libraries AI Initiative and AI Studio; the University of Rhode Island AI Lab, which is based in the URI Library; Andromeda Yelton's AI/ML-driven HAMLET (How About Machine Learning Enhanced Theses?) interface at MIT ("the first machine learning system developed by a library and deployed to production in a library anywhere in the US," according to Jason Griffey)[9]; and other AI/ML-driven projects at academic, public, and municipal libraries in the United States, United Kingdom, and Europe.[10]

Furthermore, academic and public libraries in the United States and the UK have implemented Yewno Discover, a commercial AI-driven system dedicated to "transforming information into knowledge."[11]

The Auburn University Libraries as an AI/ML Testbed

The Auburn University Libraries represent a promising testbed for artificial intelligence and machine learning projects. Established in 1856, Auburn University is a land-grant university in east-central Alabama specializing in agriculture (including fisheries and forestry), architecture, business, engineering, and the applied sciences. In early 2020, the

Auburn University Libraries formed a new Research Support Department offering an array of services in twelve areas, including research data management, grant funding and proposal development, digital scholarship and digital humanities, maximizing research impact and visibility, and IT tools and consulting. As part of this effort, the Auburn University Libraries recently constructed an Innovation & Research Commons (I&RC) on the ground floor of the main library. Among other things, the I&RC has a data services hub—the DataSpace—offering on-site expertise in using big data sets and AI/ML.

The Auburn University Libraries also joined the FOLIO Product Council in 2017 and are contributing subject matter expertise and software developer time to the project. Launched in 2016 by EBSCO Information Services, Index Data, and the Open Library Environment (OLE) consortium of academic libraries, the FOLIO Project is an ambitious international initiative to build an open source library services platform (LSP) consisting of discrete, interchangeable applications for core library functions (acquisitions, cataloging, circulation and patron management, and reporting) that can also be linked to or integrated with other content management systems (e.g., archival management systems).[12] FOLIO has shown that libraries and industry partners can build paradigm-shifting open source software applications on an enterprise scale. Combining the FOLIO core platform with AI/ML tools offers some intriguing possibilities for extending FOLIO's capabilities.

The Auburn University Libraries have had a longstanding customer relationship with EBSCO Information Services. Due to the libraries' participation in the FOLIO Project and its involvement in two National Science Foundation (NSF) Convergence Accelerator proposals, this relationship has expanded to include a research and development component, with EBSCO and the libraries working together on R&D projects of mutual interest. The NSF proposals also led to a research partnership between the Auburn University Libraries and the IBM Watson Team at the IBM Research Triangle Park Center for Advanced Studies. The partnerships with EBSCO and IBM have made it possible for the authors to explore applying AI/ML tools to current library priorities, such as developing tools and services that directly support the university's research goals.

These factors were instrumental in the decision to embark on an exploratory project using a commercial AI/ML solution. The project in question applies the IBM Watson suite of AI/ML services to the Military REACH Library. The project began in early 2020 and is still in progress at the time of writing.

The Military REACH Project and IBM Watson

As academic libraries seek new ways to align with current university research practices and to engage as vital partners in campus research activities, the Auburn University Libraries have built a successful partnership with the Auburn University College of Human Sciences (CHS) on an externally funded research initiative: the Military REACH Project.[13] The mission of the Military REACH Project is to support US military families and family readiness by translating academic research and other resources into practical applications.

This means making research both accessible to and usable by Department of Defense (DoD) family support specialists and military families themselves.

Originally based at the University of Minnesota, Military REACH moved to Auburn in late 2017 as the result of a successful proposal in response to a competitive funding solicitation from the Department of Agriculture (USDA) and the DoD. At the invitation of the project's principal investigator, Dr. Mallory Lucier-Greer of the CHS Department of Human Development and Family Studies, the libraries contributed their knowledge and expertise to the successful funding proposal that brought the Military REACH Project from Minnesota to Auburn. From the beginning of the project, the libraries' Military REACH support team has worked with their counterparts in CHS and the project's leadership to build the robust IT and bibliographic infrastructure on which the project rests. This has included setting up and configuring Military REACH servers in the Auburn University OIT Data Center; creating the Military REACH Library in DSpace; assisting in the development and hosting of the Military REACH website; providing expert guidance on metadata standards, accessibility, usage statistics, and social media; training IT support staff in CHS on the tools used in Military REACH; and helping the Military REACH team transition gradually to a new IT support structure.

Among the questions that Military REACH is designed to answer are:

- How does military deployment impact child development outcomes?
- What parenting support programs are available for military parents?
- What challenges do veterans face after leaving the military?
- What factors help service members cope with PTSD?

The source for answering these and other questions is the Military REACH Library, a DSpace repository and bibliographic database of research publications and other resources on military life and family health. It currently contains more than 3,000 documents dating from 1971 through 2020, including approximately 1,300 Military REACH Research Summaries/TRIP Reports (detailed two-page abstracts of research articles from peer-reviewed journals), as well as longer research reports on issues affecting US military families. The Military REACH Library uses a standard DSpace search interface, with filters for publication year, publication type, focus terms, military branch of service, and age group. This type of interface is not well suited for answering the kinds of natural-language questions referenced above. To make the Military REACH Library even more accessible to its intended audience, a different approach is needed.

Enter IBM Watson. Named after IBM's first CEO, Thomas J. Watson, Watson is a suite of AI/ML software tools and services built around natural language processing (NLP), natural language understanding (NLU), image and video analytics, speech recognition, and more.[14] Originally developed in the first decade of this century as part of IBM's DeepQA project, Watson gained wide public attention when it competed against human contestants on the TV quiz show *Jeopardy!*, winning the first prize of $1 million in 2011. Today, Watson is used in health care, construction, education, finance, weather forecasting, fashion design, and other areas. Watson can be used to surface concepts, categories, sentiment, and emotion, and to apply knowledge of unique entities in the subject domain to the target data. With Watson Knowledge Studio (WKS), researchers can define domain-specific ontologies and use them to enhance NLU engines for better responses. Other Watson services (e.g., text-to-speech/speech-to-text) can be used to build voice-based interfaces for AI expert systems.

The decision to use Watson for this project was inspired by the experience of working with researchers from the IBM Research Triangle Park Center for Advanced Studies on one of the NSF proposals referenced above. Although that proposal was unsuccessful, conversations with the IBM team suggested possibilities for using Watson on local repositories and datasets. The Military REACH Library was identified as an excellent candidate for an exploratory AI/ML project. The authors pitched the idea to the IBM RTP-CAS team in late 2019 and began working on it in January 2020.

The goal in the first phase of this exploratory project is to develop a Watson-driven natural-language query interface for non-copyrighted documents in the Military REACH Library. To that end, two of the services in the IBM Watson Studio, Watson Discovery and NLU, have been applied to a sample set of 1,285 PDF documents from the Military REACH DSpace repository. This involved the following operations:

Creating a Watson sandbox. Working with the Watson team at the IBM RTP-CAS in Raleigh, North Carolina, the AUL IT staff set up a complimentary Watson Discovery account in the IBM Cloud space. The account is for research purposes; access to it is currently limited to the project team members at the Auburn University Libraries and IBM RTP-CAS.

Document extraction. AUL IT staff extracted 1,285 non-copyrighted research summaries in PDF form from the Military REACH Library in DSpace and uploaded them as a research collection to the AUL Watson Discovery account in IBM Cloud (figure 12.1).

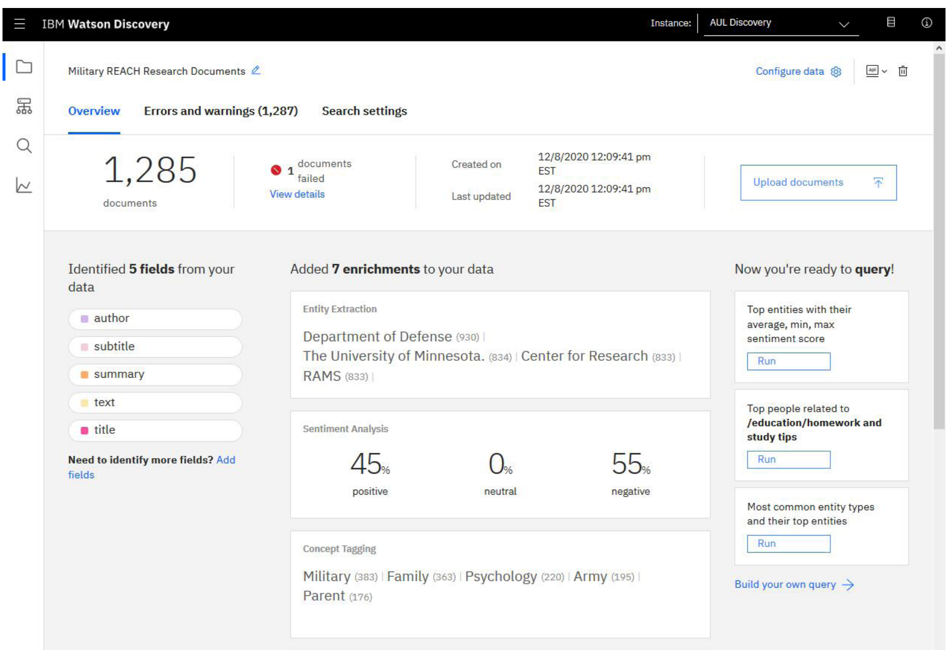


Figure 12.1
The Military Reach Research documents collection in IBM Watson Discovery.

The research summaries in the training dataset are structurally and semantically similar, with each of the automatically selected forty-four sample documents falling into one of two templates. Document sections are clearly delineated with headers and white space, which resulted in object identifications by Watson that were relatively unambiguous. This implies that such tools can more reliably identify fields and objects in otherwise unstructured data if the documents are formatted for human readability. While Watson can detect patterns corresponding to cohesive blocks of information within the document, it is up to the trainers to define which parts of the documents are of interest and how they should be represented. In this test case, the relative homogeneity of the data has allowed the authors to select field identifiers common to both document styles, which was important as the tool limits the number of custom fields they can create. Training Watson necessitates making judgments about which information should be highlighted or de-emphasized, as the tool is neutral toward which of the fields it identifies are worthy of focus. Depending on how it is trained, Watson could respond in multiple ways to the same underlying dataset, which means it is important to define the intended application before training begins.

Document markup and field identification. Using Watson Discovery's Configure Data tools, the authors applied color-coded field identifiers (for header, footer, image, text, title, author, summary, findings, implications, methods, and questions) to the forty-four Watson-selected sample documents in the Military REACH collection (figure 12.2).

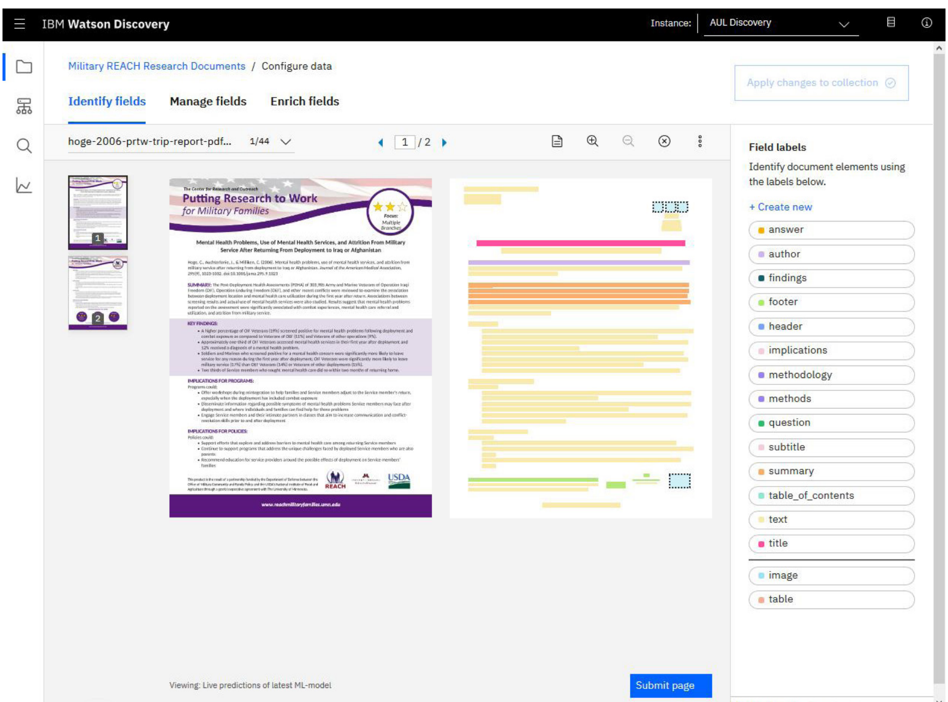


Figure 12.2

Document markup and field identification in IBM Watson Discovery.

Once the trainers have settled on a system for describing the underlying information in a useful way, the process of marking up the training documents themselves is straightforward. Field labels are selected and applied to box-shaped areas of the document that correspond to the underlying data. As the process continues with new examples, the tool begins to predict and assign labels to some objects within the training documents on its own, leaving the human trainer to correct any mistakes or apply labels to areas that were missed. For instance, if the first line of the document is marked as *title* frequently enough, the next document presented to the trainer might already have the box occurring in that area labeled as *title*, but it might not yet detect multi-line titles. In that case, the trainer would designate the box or two immediately below the first one as *title* until Watson "understands" that the end of the title is not the same as the end of the first line, but the beginning of the white space between sections. This type of work does not require extensive technical knowledge on the part of the trainer. It is important, however, for the trainer to understand what should be included and excluded from each field they are asked to assign.

Collection enrichment. Watson Discovery automatically enriches (adds cognitive metadata in JSON to) the text field in ingested documents with semantic information collected by four Watson NLU enrichment functions: entity extraction, sentiment analysis, category classification, and concept tagging. Using Watson Discovery's Configure Data tools, the authors added four more enrichment functions to the test collection: keyword extraction, relation extraction, emotion analysis, and semantic role extraction (figure 12.3).

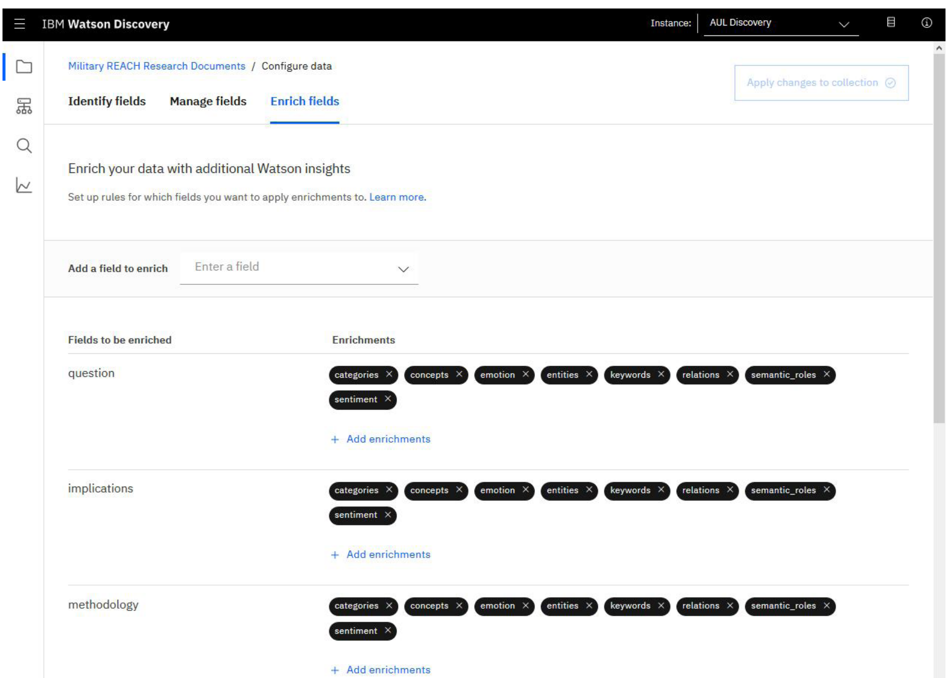


Figure 12.3

The Field Enrichment Tool in IBM Watson Discovery.

Building natural language queries. This is where the project stands at the time of writing. The authors are working with the principal investigator and subject matter experts on the Military REACH Project to identify natural language questions in this area and run them against the test collection to see whether the results are plausible or whether they fall into the category of what Andromeda Yelton has called "attractive nonsense."[15] Preliminary results using sample queries suggest that the documents that Watson is surfacing are clearly responsive to the test queries (figure 12.4).

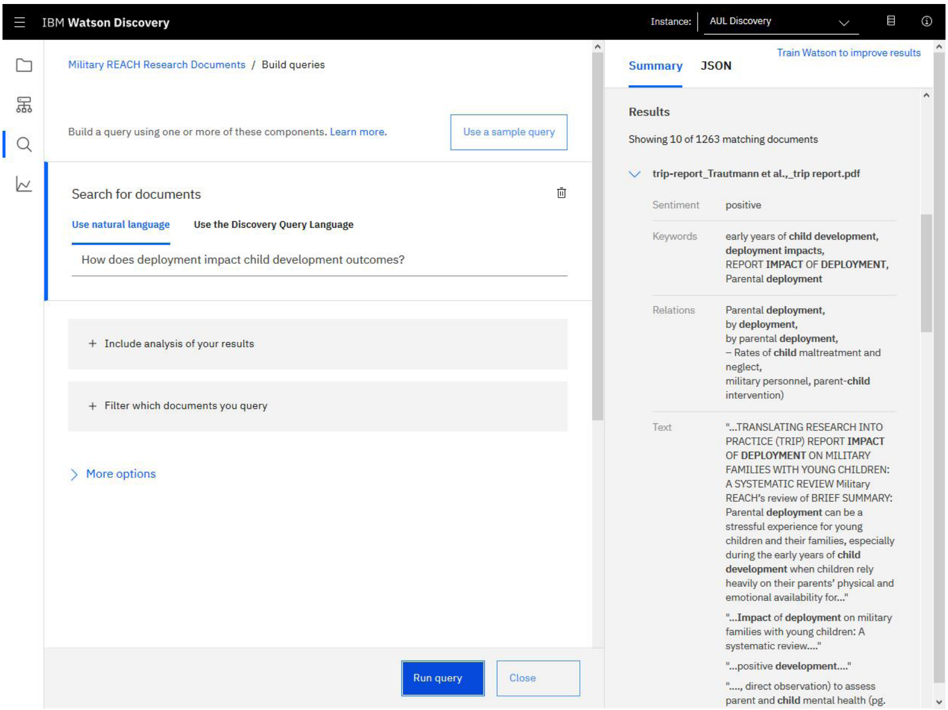


Figure 12.4
Preliminary results of a natural language search query in IBM Watson Discovery.

Next steps in this project include the following:

Adding more materials to the test collection. At 1,285 documents, the Military REACH test collection is too small to allow a proper assessment of Watson's capabilities. The authors have been using it primarily to become familiar with Watson's features and to identify stumbling blocks. An insufficiently large body of material is one of Yelton's "traps for the unwary."[16] This can be remedied by including documents from the Auburn University institutional repository (AUrora) and database of electronic theses and dissertations (AUETD). Taken together, these sources should add approximately ten thousand items to the test collection. Eventually, the authors intend to add materials harvested from the Public Library of Science (PLoS) and relevant public datasets from Data.gov.

Building a public-facing natural language query interface and recommendation engine. The user interface and content display are far from being ready for unveiling as an alpha version, much less a production version. The IT team is working with the

IBM Watson team on this, using APIs that have been developed for other collections. In addition, the Computer Science and Software Department in the Samuel Ginn College of Engineering at Auburn University is building a Senior Design Project around using IBM Watson Assistant to make the Military REACH Library more accessible to users. The same department is also working with the Military REACH team on using AI tools and techniques to search the scholarly literature for relevant publications and presentations for inclusion in the Military REACH Library. The authors hope to unveil an alpha version of the natural-language query interface in the second half of 2021. The long-term goal is to use IBM's speech-to-text/text-to-speech services to develop an AI-driven chatbot that service providers and military family members can access on their mobile devices or personal digital assistants.

Enlisting Military REACH and other subject matter experts in this field in training the system using IBM Knowledge Studio. This will be an iterative process as the authors add more material to the Military REACH test collection. Watson Discovery has a built-in training module that we intend to use for this purpose (figure 12.5).

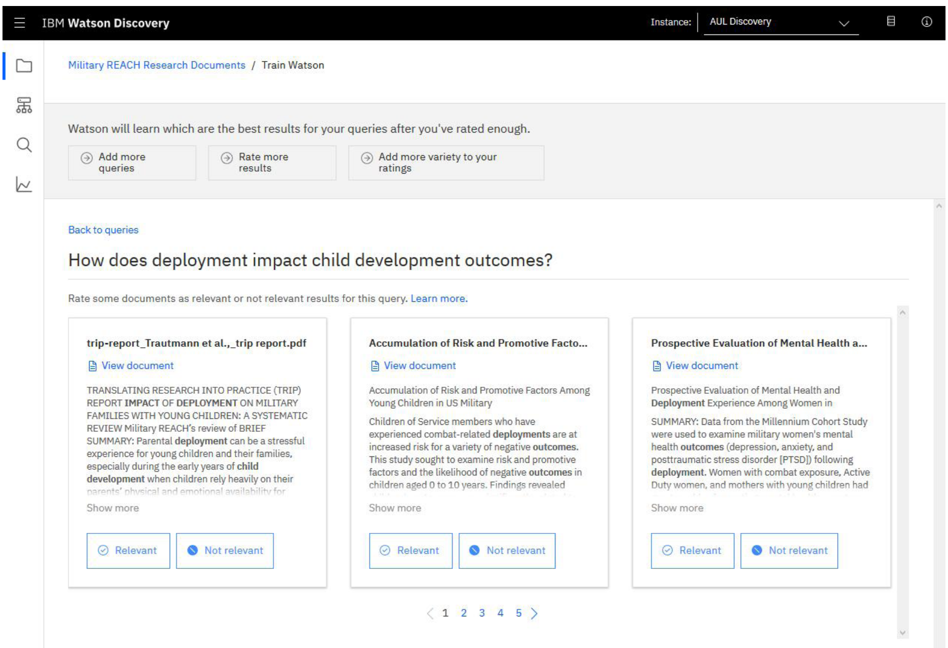


Figure 12.5
The IBM Watson Discovery training interface.

Lessons Learned

This project has been extremely instructive in highlighting the possibilities and challenges of AI/ML in academic libraries. Perhaps counterintuitively, the authors discovered that implementing AI/ML solutions is an extremely labor-intensive task. AI/ML tools still require extensive human intervention—for example, in the areas of document markup and system training—in order to be effective. The human factor is still essential, which should be encouraging news for librarians who are worried about being replaced by AI-driven bots. Alternatively, it means that some, perhaps many, library jobs may become redundant or irrelevant, and currently employed librarians will have to learn a challenging set of new skills.

AI/ML solutions and tool suites are extremely complex and are often not turnkey solutions. Major DIY assembly is required in the form of IT support, system administration, collection markup and enrichment, training, and interface design. Libraries interested in working with AI/ML tools will need access to an IT department with a deep and varied skill set, either locally or at the institution level. Getting buy-in will be difficult or impossible if AI/ML projects do not connect directly to library or institutional priorities. Even if there is support for AI/ML projects, creating the necessary IT bandwidth will almost certainly be a challenge. Campus IT is also likely to be involved.

Another challenge for academic libraries is cost. Commercial AI/ML solutions are expensive. The tiered price schedule for a production-grade instance of IBM Watson Discovery, just one service in the Watson suite of services, starts at $500 per month.[17] The Auburn University Libraries were fortunate in having a well-established working relationship with IBM that allowed them to negotiate complimentary R&D platforms for the project. The authors were also fortunate in having a project that aligned with the company's product development plans. The authors expect that the terms of the conversation will change once they move into the production phase. Establishing mutually beneficial R&D relationships with vendors at the outset may enable libraries to negotiate more favorable deals for jointly developed platforms.

As with any technology, there is a concern about the potential for misuse and misinterpretation of results. The process of training AI/ML algorithms necessarily involves allowing them to process pieces of information to find emergent patterns that reflect the machine's understanding rather than a human's understanding. Librarians who help to develop and support such applications should be mindful to supply additional context when needed to assist users in understanding the proper scope of the tools and avoiding unwarranted leaps of logic.

Despite the challenges noted above, there are upsides to working on AI/ML projects. There is nothing like learning by doing and applying industry-strength AI/ML tools on real collections for real purposes. This was far more instructive and useful than learning about them in the abstract or in a classroom situation. The authors have a clearer understanding of how AI/ML tools work, what they can realistically be used for, what to watch out for (Yelton's "traps for the unwary"), and areas where they, and the use of them, can be improved.

This experience has equipped the authors to understand "how and what an outside vendor could be doing in the training stages" and to help ensure that commercial AI/ML products meet the needs of libraries.[18] In other words, it has reduced the *black box* factor, the opacity of AI/ML systems, that Griffey, Yelton, and other observers have identified. Trust is good, but informed, hands-on experience is better. Suggestions have already been made to the IBM Watson team for developing a more librarian- and user-friendly interface for Watson Discovery.

Conclusion

Despite missteps and missed targets, in the project's original timeframe, progress was made toward functional prototypes of a Watson-driven interface for the Military REACH Library. This project has already helped point the way for future development, enhancements, research collaborations, and (possibly) research funding proposals.

The experience with Watson has positioned the authors to help other researchers at Auburn explore the use of AI/ML tools in their own work, including the conceptualization and drafting of external funding proposals. This is directly in line with the university's and the libraries' research priorities and strategic goals.

Acknowledgments

The authors would like to thank Clint Bellanger (Auburn), James Jushchuk (IBM), Dr. Mallory Lucier-Greer (Auburn), Anand Singh (IBM), and Michael Stone (Auburn) for their help with the chapter and the projects described therein.

Endnotes

1. Peter Fernandez, "'Through the Looking Glass: Envisioning New Libraries Technologies' How Artificial Intelligence Will Impact Libraries," *Library Hi Tech News* 33, no. 5 (2016): 5–7.
2. Andrew M. Cox, Stephen Pinfield, and Sophie Rutter, "The Intelligent Library: Thought Leaders' Views on the Likely Impact of Artificial Intelligence on Academic Libraries," *Library Hi Tech* 37, no. 3 (2019): 418–35.
3. Cox, Pinfield, and Rutter, "The Intelligent Library," 431.
4. Carl Benedikt Frey and Michael A. Osborne, "The Future of Employment: How Susceptible Are Jobs to Computerisation?," *Technological Forecasting and Social Change* 114 (January 2017): 254–80.
5. Kenning Arlitsch and Bruce Newell, "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries," *Journal of Library Administration* 57, no. 7 (2017): 796.
6. Barbara A. Wood and David J. Evans, "Librarians' Perceptions of Artificial Intelligence and Its Potential Impact on the Profession," *Computers in Libraries* 38, no. 1 (January/February 2018): 26–30.
7. Sarah Myers West, Meredith Whittaker, and Kate Crawford, *Discriminating Systems: Gender, Race and Power in AI* (New York: AI Now Institute, 2019), 10.
8. "RLI 299: Ethics of Artificial Intelligence," Association of Research Libraries, *Research Library Issues*, 2019.
9. Andromeda Yelton, "HAMLET: How About Machine Learning Enhancing Theses?," https://hamlet.andromedayelton.com/; Jason Griffey, "Introduction," in *Artificial Intelligence and Machine Learning in Libraries*, ed. Jason Griffey (Chicago: American Library Association, 2019), 9.
10. Pirjo Kangas, "AI Use Cases," AI for Librarians, https://www.aiforlibrarians.com/ai-cases/; "Projects in Artificial Intelligence Registry (PAIR): A Registry for AI Projects in Higher Ed," University of Oklahoma, https://pair.libraries.ou.edu/search.
11. "Yewno Discover," Yewno, https://www.yewno.com/discover; see also Library Technology Guides, https://librarytechnology.org/news/searchresults/?SID=2020062833630371&code=&code=BIB&Company=Yewno&sort=chron.
12. FOLIO Project, https://www.folio.org/.
13. Military REACH Project, https://militaryreach.auburn.edu/.
14. "IBM Watson products and solutions," IBM, https://www.ibm.com/watson/products-services.
15. Andromeda Yelton, "Chapter 2: HAMLET: Neural-Net-Powered Prototypes for Library Discovery," in *Artificial Intelligence and Machine Learning in Libraries*, ed. Jason Griffey (Chicago: American Library Association, 2019), 15.
16. Yelton, "Chapter 2: HAMLET: Neural-Net-Powered Prototypes," 13–14.

17. "Watson Discovery: Pricing," IBM, https://www.ibm.com/cloud/watson-discovery/pricing. There is a free "Lite" option, but it is not suitable for development or production.
18. Griffey, "Introduction," 8.

Bibliography

Arlitsch, Kenning, and Bruce Newell. "Thriving in the Age of Accelerations: A Brief Look at the Societal Effects of Artificial Intelligence and the Opportunities for Libraries." *Journal of Library Administration* 57, no. 7 (2017): 789–98. https://www.tandfonline.com/doi/full/10.1080/01930826.2017.1362912.

Association of Research Libraries. "RLI 299: Ethics of Artificial Intelligence." *Research Library Issues*, 2019. https://publications.arl.org/rli299/.

AUrora (Auburn University IR). http://aurora.auburn.edu/.

Cox, Andrew M., Stephen Pinfield, and Sophie Rutter. "The Intelligent Library: Thought Leaders' Views on the Likely Impact of Artificial Intelligence on Academic Libraries." *Library Hi Tech* 37, no. 3 (2019): 418–35. https://doi.org/10.1108/LHT-08-2018-0105.

———. "'Through the Looking Glass: Envisioning New Libraries Technologies' How Artificial Intelligence Will Impact Libraries." *Library Hi Tech News* 33, no. 5 (2016): 5–8. https://www.emerald.com/insight/content/doi/10.1108/LHTN-05-2016-0024/full/html.

FOLIO Project. https://www.folio.org/.

Frey, Carl Benedikt, and Michael A. Osborne. "The Future of Employment: How Susceptible Are Jobs to Computerisation?" *Technological Forecasting and Social Change* 114 (January 2017): 254–80. https://doi.org/10.1016/j.techfore.2016.08.019.

Griffey, Jason, ed. "Artificial Intelligence and Machine Learning in Libraries." *Library Technology Reports* 55, no. 1 (January 2019). https://doi.org/10.5860/ltr.55n1.

IBM. "IBM Watson products and solutions." https://www.ibm.com/watson/products-services.

Kangas, Pirjo. *AI for Librarians*. https://www.aiforlibrarians.com/.

Keller, Michael, and Ruth Pickering. "Part 2: AI in the Research Library Environment." *Charleston Library Conference: Artificial Intelligence in Scholarly Research* (Webinar), May 20, 2020. https://charlestonlibraryconference.com/artificial-intelligence-in-scholarly-research-part-2/.

Military REACH Project. https://militaryreach.auburn.edu/.

National Science Foundation. "Artificial Intelligence (AI) at NSF." https://www.nsf.gov/cise/ai.jsp.

———. "NSF's 10 Big Ideas: Harnessing the Data Revolution." https://www.nsf.gov/news/special_reports/big_ideas/harnessing.jsp.

———. "NSF Convergence Accelerator (C-Accel)." https://www.nsf.gov/od/oia/convergence-accelerator/index.jsp.

Stanford University Libraries. "Artificial Intelligence." https://library.stanford.edu/projects/artificial-intelligence.

University of Calgary. Libraries and Cultural Resources. "Critical Roles for Libraries in Today's Research Enterprise." https://library.ucalgary.ca/libraryresearchplatform/symposium.

University of Oklahoma. University Libraries. "Projects in Artificial Intelligence Registry (PAIR): A Registry for AI Projects in Higher Ed." https://pair.libraries.ou.edu/search.

University of Rhode Island. "AI Lab." https://web.uri.edu/ai/.

West, Sarah Myers, Meredith Whittaker, and Kate Crawford. *Discriminating Systems: Gender, Race and Power in AI*. New York: AI Now Institute, 2019. https://ainowinstitute.org/discriminatingsystems.html.

Wood, Barbara A., and David J. Evans. "Librarians' Perceptions of Artificial Intelligence and Its Potential Impact on the Profession." *Computers in Libraries* 38, no. 1 (January/February 2018): 26–30. http://www.infotoday.com/cilmag/jan18/Wood-Evans--Librarians-Perceptions-of-Artificial-Intelligence.shtml.

Yelton, Andromeda. "HAMLET: How About Machine Learning Enhancing Theses?" https://hamlet.andromedayelton.com/.

———. "Chapter 2: HAMLET: Neural-Net-Powered Prototypes for Library Discovery." In *Artificial Intelligence and Machine Learning in Libraries*, ed. Jason Griffey, 10–15. Chicago: American Library Association, 2019.

Yewno Discover. https://www.yewno.com/discover.

Part III
Toward Future Applications

Chapter 13

Ethical Implications of Implicit Bias in AI:

Impact for Academic Libraries

Kim Paula Nayyer and Marcelo Rodriguez

Introduction

Academic libraries are exploring artificial intelligence (AI) applications that have the potential to create new or improved user experiences, streamline ways of working, and deliver new insights to their activities. Nevertheless, it is now clear that AI applications are not neutral technological solutions. They can embed and magnify prejudices and stereotypes, and they can perpetuate errors and limitations in training and accumulated datasets. At the same time, academic libraries abide by ethical considerations of social responsibility. If datasets and algorithmic *black boxes* in AI systems replicate or aggravate inappropriate discrimination in their use of information, or if they simply lack or ignore data, they can produce distorted outcomes. The ethical implications for academic libraries and end-users can be profound.

This chapter examines these issues, illustrates problematic outcomes, and identifies both the need for caution and some paths to the ethical use of AI applications in academic libraries. After a general exploration of the essence of machine learning (ML), this chapter explains what implicit bias is, how it enters ML applications, and why the problem is insidious and challenging. The authors present an illustrative review of the ethical foundations of the work of academic libraries and draw analogies to other professional interfaces with AI and implicit bias. Possible scenarios of ethically problematic outcomes in academic libraries are explored.

What Is Implicit Bias and How Does It Enter AI Applications?

To understand how AI applications can embody implicit bias, one must understand how modern AI works. Until a few decades ago, the phrase artificial intelligence conjured images of futuristic technologies ranging from friendly robotic companions to mission-fulfilling supreme computers like HAL 9000 of the classic film and book *2001: A Space Odyssey*.[1] Now, rather than being fanciful or theoretical, AI applications are present everywhere in modern life. Smartphones learn to spell friends' and family members' names, streaming video services seem to know what viewers want to watch before they do, and popular search engines predict what users are searching for after a few keystrokes.

The advent of powerful computing and vast amounts of usable data in digital form continue to push the proliferation of ML applications, and they are now commonplace in environments in which people provide and use services.[2] Generally, in library environments and elsewhere, AI is seen not as omniscient standalone operators but rather as components of larger products and processes.[3] AI tools perform specific tasks in service of the goals of the system. To a greater or lesser degree and with varied effectiveness, the AI component replicates, or at least replaces, human thought in the fulfillment of that task.

Machine Learning: An Algorithm, Its Training Data, Iterative Learning, and Algorithm Self-Revision

Early AI tools were developed using an expert-systems approach sometimes called "Good Old-Fashioned AI" or GOFAI.[4] The algorithm is a complex decision tree, coded to draw from the deep knowledge of human experts. The algorithm is fed information it needs to make decisions that would replicate the decisions of someone with that realm of knowledge.[5] Any errors, by miscalculation, from bias or otherwise, derive only from the decision instructions and rules the algorithm is given. The GOFAI tools made calculated decisions, but they did not learn. Their knowledge was static, based on the expertise and rules supplied to them.

Conversely, the powerful AI tools and processes of today and the near future have machine learning at their core. The machine learning-driven AI algorithm differs from simple executable algorithms. It does not merely draw from an encyclopedic body of knowledge and execute a task its algorithm instructs it to do. Rather, the ML algorithm is written to cause the tool to train itself from data supplied to it and to learn from its own executions. The learning is iterative: the machine learns first from training datasets to perform its function and usually to make simple decisions of a predictive nature. These decisions follow outcomes in the dataset the tool is fed and then add to the content of the dataset. The machine learns from its previous predictions and activity grows its data. From a large set of data and outcomes, the machine predicts what may be similar to previous choices.

The Implicit Bias Problem: Humans, Data, and Bias

The term "bias" is associated with prejudgment, reliance on irrelevant factors, negatively imbalanced outcomes, and often connotes a conscious exercise of moral unfairness. For example, any bias in the output of a GOFAI tool closely derives from bias in its algorithm or the expert knowledge in its encoding. This kind of explicit bias certainly occurs and can infect a machine learning process, but it isn't the whole story.[6] Biases can also enter implicitly into ML as humans subconsciously form schemas or put things in categories to manage vast amounts of information. This process may help organize the environment, but it can also have more serious impacts. A variable can come to represent a class that shares certain characteristics; for example, someone may inappropriately ascribe to a single individual a host of other features sometimes associated with that class. Implicit bias can cause superficially or facially neutral factors to produce imbalanced outcomes. This kind of bias can enter implicitly into ML applications and can give rise to subtle but problematic outcomes.[7]

Machine learning algorithms learn from data and coding produced by flawed human systems, data that derives from past bad decisions, and data that derives from systemic and societal injustices.[8] Sources of implicit bias in AI tools are multifold, and the impact of these sources are not equal. Neither are they all equally well understood. Some are identifiable and, through some work and attention, can be addressed, while others are subtle and more difficult to resolve: bias can enter an ML process insidiously, through pervasive and deeply ingrained systems that govern our societies.

Some literature on AI and libraries points to human coders as a key source of bias. The suggestion is that human biases imbue biased algorithms and are thus embedded into AI tools via those algorithms.[9] They assert that the source of AI bias is human error or the biased human's role in writing an algorithm.[10] To be sure, subconscious and conscious biases accompany everything humans do; indeed, early AI proponents cited this human fallibility as a reason to let the machines make the decisions.[11] The bulk of the implicit bias is the data from which the algorithm learns and with which it operates. To grasp this requires an understanding of how current AI tools work.

A human may identify a problem, and then other humans proceed to design and code an algorithm to address that problem. The algorithm is trained to solve this by generating output based on an initial set of data. That dataset allows the algorithm to generate an analysis and a prediction of an outcome based on coded or labelled data and statistical probabilities. The outcome of the analysis supplements the dataset, and the algorithm continues to learn and change. The trained algorithm can then be applied to larger datasets, executing predictions from the new data, and continuing to learn or train itself from this data. The process continues iteratively, with more data added and the algorithm continuing to learn and change.[12]

Unsupervised Machine Learning and Supervised Machine Learning

In unsupervised machine learning, the algorithm runs on the dataset and produces its predictions or outputs on its own. There is no human intervention to reduce distortion or to evaluate results during the iterative learning process. This means any problematic results in the outputs can continue to feed the iterative process. If, for example, a dataset used to train a tool to predict challenging library patrons draws from big datasets of national criminal justice data, the tool's output likely will reflect historical societal and racial inequities that do not appropriately predict for the intended question. When this data is used to train a tool, serious implicit bias in that data can go undetected, all while the algorithm is adapting and refining itself from that data. The result is not only a replication but also an amplification of the initial biases in the data. The multiple iterations of the ML process cause the problems with the algorithm to become difficult to even identify, let alone resolve.

Supervised machine learning, on the other hand, incorporates some measure of human intervention and, along with it, a more ethics-driven approach. A person who is familiar with the dataset itself and the problem-solving algorithm, as well as with desired or predicted outcomes, will audit the dataset and the algorithm's initial outputs. They will assess likely accuracy and appropriateness of the data and will select and remove features with a goal to minimize distortions. The person will look for patterns and features the algorithm initially captured and will study the relationship between them and the initial outcome. With reference to an expected outcome, the supervisor will determine which variables to include and which to exclude in the model.

Ethical Foundations of Academic Libraries: Guidance for Addressing AI and Bias Issues

Over the last two decades, as AI has begun to see practical use cases, ethical guidelines seem to have proliferated. Universities, faculty, librarians, professional organizations, companies, governments, international organizations, and civil society all strive to provide frameworks, principles, or statements to guide solutions in their sectors. Overlaps even exist among the multiplicity of individual and localized efforts. Whereas academic libraries do not yet have explicit guidance for addressing ethical issues arising from AI, they can use both their existing framework and guidance from other sectors as resources to guide decisions and to support concrete steps—and perhaps to guide the creation of a statement. Further, given the multitude of actors who interact with academic libraries—faculty, students, vendors, government, the general public, or others—benefits can be drawn from awareness of ethical considerations that might guide other entities.

Academic Libraries and the ACRL Framework

An exploration of ethical guidance for academic libraries must begin with the *Framework for Information Literacy for Higher Education*, adopted in 2016 by the Board of the Association of College & Research Libraries (ACRL).[13] The ACRL *Framework*'s six concepts represent core values for higher education institutions to guide their teaching and researching tasks. The ACRL *Framework* strives to help academic institutions frame their own mission statements, scholar goals, and outcomes under shared concepts and ideas. Its conceptual understandings aim to provide the philosophical stepping stones that can assist academic institutions to develop their own tools, policies, and benchmarks for teaching and research.

Despite no explicit appearance of the word *ethics* in the text of the *Framework*, a few of its six concepts may help academic libraries construct their own ethical considerations. Examples include Authority is Constructed and Contextual, Information Creation as a Process, Information Has Value, and Searching as Strategic Exploration. Taken together, these concepts and their underlying statements set out exactly what the *Framework* was intended to facilitate: an initial set of steps for academic libraries to frame their own approaches to challenges, both current and unforeseen.

The *Framework* encourages academic libraries to challenge the entire information environment. It challenges embedded cultural, social, and political assumptions and the biases and context that underlie data, self-described evidence, research questions, methodologies, and conclusions. As the uses of the *Framework* continue to evolve, academic libraries are prompted to ask questions and pursue answers. The imminent applications of AI in academic libraries call for the *Framework* to be more explicit regarding ethics and bias in AI.

Universities and Ethical AI Imperatives

Efforts in universities are relevant to academic libraries not only for the academic institutional environment but also because university research activities can launch AI tools useful for academic libraries and give guidance for their ethical, bias-free application. Since the early 2010s, university scholars have urged robotics and robotics engineers to develop codes of ethics, and ethical guidance for the initial research stages is a continued imperative.[14] Some universities have created their own recommendations that build upon their own histories as well as current AI thought leadership, and they may serve as models for other universities.[15] Universities also are at the forefront of conversations about AI ethics through research centers and think tanks dedicated to this issue. A few examples highlight the development of spaces for consideration of ethical uses of AI, and they also illustrate initiatives that strive to bring numerous scholars together under the umbrella of multidisciplinary approaches to ethics in AI applications. For example, Stanford University collaborated with the Machine Intelligence Research Institute (MIRI) in 2006. MIRI's conferences and national events set the stage for conversations exploring concepts such as "friendly AI" and "effective altruism." In 2017, the MIT Media Lab and the Harvard Berkman-Klein Center for Internet and Society launched the Ethics and

Governance of AI Initiative. This initiative strives to both build a national network of universities, companies, and civic organizations and to financially support AI research and projects for the public interest. In the same year, New York University founded the AI Now Institute. Through its advocacy, research, symposia, and expert testimony, AI Now has positioned itself as a key actor in a national and international conversation of social implications and applications of AI research. More recently, in 2019, Stanford University launched the Institute for Human-Centered Artificial Intelligence (HAI), intended to be a cross-disciplinary hub for all faculties within Stanford University and other universities and colleges.

AI, Implicit Bias, and Ethical Responsibilities and Opportunities for Academic Libraries

Responsibilities of AI Innovators in Academic Libraries

Component tasks that are the responsibility of AI technologies can serve different types of functions. Some AI functions are in use in libraries today, mainly in predictive or decision-support contexts, and more extended applications in the future are foreseen. By their iterative and cumulative learning nature, the developed operation of machine learning algorithms and the basis for the outcomes they produce are often unknown to even the initial coder.[16] Even if inadvertently, developers are likely creating AI tools that magnify systemic biases and distort outcomes. Unfortunately, many innovators may not be alert to the reality that implicit bias becomes embedded this way. They may not be familiar with the extent to which implicit bias in datasets can drive ML or how readily and regularly they exacerbate historical patterns of discrimination. Innovators using ML algorithm programming skills may build exciting tools to carry out valuable functions in libraries. However, if innovators use weak training data tainted by bias, unsupervised learning processes, or even supervised learning processes without sufficient understanding of variables and appropriate outcomes, they can unwittingly produce outcomes tainted by and perpetuating implicit bias.

Bias Mitigation in Supervised Machine Learning Can Intercede and Minimize Errors

Researchers are working to create methods to mitigate bias.[17] They are able to supply corrective data and decision-making paths to help train AI tools in different ways. They may work to fix datasets and alleviate concerns regarding embedded implicit bias. These processes may require extensive testing and repeated interventions to give statistically balanced results. They also require knowledgeable individuals who have, in addition to ML expertise, a good understanding of potential problems and expected outcomes. Some working in this area also advocate for regulatory guidance to provide a framework for AI.[18]

A related concern is that the power to develop AI tools relates directly to access to, ownership of, or control over big datasets.[19] Most amassed big data that could be useful to academic libraries and information workers are held by a few stakeholders, and the masses are generally unaware of what is in that data or the details of the algorithms. Developers themselves, privately and within academic libraries, may never fully understand the nuances of AI-driven library solutions which, by their self-developing nature, become unknown to either the original human coder or the user. Ethical adoption of AI-driven solutions requires academic libraries to fully evaluate this reality.

A challenge to the development of locally created AI applications for use in academic libraries is likely to be a shortfall in large, representative amounts of high-quality usable data pertinent to the local community and the purpose. Many sources of openly available training datasets exist, but these will only assist if the focus of the data and their presentation are suitable to the desired machine learning goal and serve the project.[20] Even then, there are no assurances that the available data have not been drawn from sources that might have been subject to decades of systemic, institutional, and unconscious (or even explicit) bias.

We know that uncertainty exists in our ability to assess the fairness and reliability of AI tool outputs. Users without an ability to break down the results from an AI *black box* may have continued uncertainty about whether a prediction, a recommendation, or a decision is fair or embeds and perpetuates implicit biases. This raises the concern of whether use of the tool is a violation of academic library professional standards, patron respect, or ethical standards.

Supervision allows a human to step back and assess the value of the variables and the propriety of including any variable in the ML process. However, the availability of expertise can be a challenge: an effective supervisor must have broad and deep knowledge of the issue and model or problem-solving algorithm, the range of variables, and likely predicted outcomes. Academic librarians can be ideal partners for developers searching for clean datasets for their models. Because of the contribution of their longstanding and novel expertise, academic libraries are well-positioned to explore, experiment, and work with researchers, scholars, and practitioners to provide meaningful, creative, and ethical solutions to the problems of AI and implicit bias.

Conclusion

As machine learning technologies advance and data sources become more unwieldy and opaque, concerns about embedded and inextricable implicit bias are both real and increasingly widespread. At the same time, computing power and the abundance of datasets further the proliferation of AI applications that can be used in or by academic libraries.

Simultaneously, communities of practice and professions are advancing valuable ethical assessments that can assist in clarification and guidance for academic libraries.

As centers of information, knowledge, and creation, academic libraries can play a pivotal role in clarifying issues with data, AI, and implicit bias. The ethical implications of implicit bias in the creation and use of AI in academic libraries is an ongoing conversation, and a range of disciplines and communities offer valuable ethical guides. These resources can help inform a framework that is current, relevant, and robust in the face of AI challenges. Academic libraries can be increasingly adept at ethically addressing issues of AI and implicit bias for their own work and that of the wider AI development community.

Endnotes

1. David G. Stork, "The Best-Informed Dream: Hal and the Vision of 2001," in *Hal's Legacy: 2001's Computer as Dream and Reality*, ed. David Stork (Cambridge, MA: MIT Press, 1997), 7–15; HAL 9000 was the name of the spaceship computer in the film and book. Interestingly, the fictional HAL 9000 of 1968 bore the power—and ultimately the weight—of ethical decision-making. HAL 9000 is often seen as the antagonist or villain of the story, as he determines the way to adhere to his programmed rules is to kill his human space travel companions; see Daniel C. Dennett, "When HAL Kills, Who's to Blame? Computer Ethics," in *Hal's Legacy: 2001's Computer as Dream and Reality*, ed. David Stork (Cambridge, MA: MIT Press, 1997), 203–14.
2. Elizabeth A. Waraksa, ed., "Research Library Issues. RLI 299: Ethics of Artificial Intelligence," Association of Research Libraries (2019): 5.
3. AI of fiction and dystopian futures is often generative AI, whose development is not yet within modern technological reach.
4. Charles Morgan, "Responsible AI: A Global Policy Framework," *ITechLaw* (November 18, 2019): 20, https://www.itechlaw.org/ResponsibleAI.
5. Morgan, "Responsible AI," 20–21; text accompanying endnotes, 24–29.
6. Ayanna Howard and Jason Borenstein, "The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity," *Science and Engineering Ethics* 24, no. 5 (2018): 1521–36, https://doi.org/10.1007/s11948-017-9975-2.
7. And sometimes biases such as stereotyping produce a nefarious circumstance where a decision-maker renders a biased judgment when fully aware of and intending to exercise that bias. This kind of explicit bias is detrimental and no doubt occurs, although it is not the focus of this chapter. For a thorough study of this, see Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York, NYU Press, 2018), http://ebookcentral.proquest.com/lib/cornell/detail.action?docID=4834260.
8. Noble, *Algorithms of Oppression*; a good encapsulation of the problem is presented by Joy Buolamwini, founder of the Algorithmic Justice League: Joy Buolamwini, "PBS NewsHour: America in Black and Blue 2020: PBS NewsHour Weekend Special," WCNY Video, 2020, https://video.wcny.org/video/america-in-black-blue-2020-a-pbs-newshour-weekend-special-4lpnmp/; see also Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Broadway Books, 2016), http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1109940&site=ehost-live.
9. See, e.g., Ellyssa Kroski and Robert J. Ambrogi, "The Future of AI in Law Libraries," in *Law Librarianship in the Age of AI* (Chicago, IL: ALA Editions, 2020).
10. See, e.g., Jamie J. Baker and Robert J. Ambrogi, "AI in Legal Research," in *Law Librarianship in the Age of AI* (Chicago, IL: ALA Editions, 2020).
11. See, e.g., John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955," *AI Magazine* 27, no. 4 (2006): 12, http://www-formal.stanford.edu/jmc/history/dartmouth.pdf.
12. A human, no doubt, usually will select an initial dataset and be involved in labelling but usually does not have control of each data point.

13. *Framework for Information Literacy for Higher Education*, American Library Association, February 9, 2015, accessed September 16, 2020, http://www.ala.org/acrl/standards/ilframework.
14. Illah Reza Nourbakhsh, *Robot Futures* (Cambridge, MA: MIT Press, 2013); Brandon Ingram, Daniel Jones, Andrew Lewis, Matthew Richards, Charles Rich, and Lance Schachterle, "A Code of Ethics for Robotics Engineers," in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (IEEE, 2010), 103–104.
15. Julie A. Shah and Melissa Nobles, "Working Group on Social Implications and Responsibilities of Computing Final Report" (Cambridge, MA: MIT Schwarzman College of Computing Task Force, 2019), 1–22, http://web.mit.edu/comptfreport/sirc.pdf.
16. Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Cambridge: Harvard University Press, 2015), 320, https://doi.org/10.4159/harvard.9780674736061; the issue in this context is is summarized by Ellyssa Kroski and James M. Donovan, "Benefits, Drawbacks, and Risks of AI," in *Law Librarianship in the Age of AI* (Chicago, IL: ALA Editions, 2020).
17. "AI Fairness 360," IBM Research Trusted AI, 2020, http://aif360.mybluemix.net/.
18. Laila Paszti, "Did the OPC Get It Right? A Proposed Approach for Regulating Artificial Intelligence," Canada | Global law firm | Norton Rose Fulbright, 2020, https://www.nortonrosefulbright.com/en-ca/knowledge/publications/33076187/did-the-opc-get-it-right-a-proposed-approach-for-regulating-artificial-intelligence.
19. Pasquale, *The Black Box Society*.
20. Joaquin Vanschoren, "Democratizing Machine Learning," OpenML, 2020, https://www.openml.org/; "Public Datasets | Google Cloud," Google Cloud Public Datasets, Google (Google, 2020), https://cloud.google.com/public-datasets/; "Open Datasets: Microsoft Azure," Microsoft Azure, 2020, https://azure.microsoft.com/en-us/services/open-datasets/.

Bibliography

American Library Association. *Framework for Information Literacy for Higher Education*. February 9, 2015. Accessed September 16, 2020. http://www.ala.org/acrl/standards/ilframework.

Baker, Jamie J., and Robert J. Ambrogi. "AI in Legal Research." In *Law Librarianship in the Age of AI*. Chicago, IL: ALA Editions, 2020.

Buolamwini, Joy. "PBS NewsHour: America in Black and Blue 2020: PBS NewsHour Weekend Special." WCNY Video. 2020. https://video.wcny.org/video/america-in-black-blue-2020-a-pbs-newshour-weekend-special-4lpnmp/.

Dennett, Daniel C. "When HAL Kills, Who's to Blame? Computer Ethics." In *Hal's Legacy: 2001's Computer as Dream and Reality*, edited by David Stork, 203–14. Cambridge, MA: MIT Press, 1997.

Google. "Public Datasets | Google Cloud." Google Cloud Public Datasets. 2020. https://cloud.google.com/public-datasets/.

Howard, Ayanna, and Jason Borenstein. "The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity." *Science and Engineering Ethics* 24, no. 5 (2018): 1521–36. https://doi.org/10.1007/s11948-017-9975-2.

IBM Research Trusted AI. "AI Fairness 360." 2020. http://aif360.mybluemix.net/.

Ingram, Brandon, Daniel Jones, Andrew Lewis, Matthew Richards, Charles Rich, and Lance Schachterle. "A Code of Ethics for Robotics Engineers." In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2010, 103–104.

Kroski, Ellyssa, and Robert J. Ambrogi. "The Future of AI in Law Libraries." In *Law Librarianship in the Age of AI*. Chicago, IL: ALA Editions, 2020.

Kroski, Ellyssa, and James M. Donovan. "Benefits, Drawbacks, and Risks of AI." In *Law Librarianship in the Age of AI*. Chicago, IL: ALA Editions, 2020.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955." *AI Magazine* 27, no. 4 (2006): 12. http://www-formal.stanford.edu/jmc/history/dartmouth.pdf.

Microsoft Azure. "Open Datasets: Microsoft Azure." 2020. https://azure.microsoft.com/en-us/services/open-datasets/.

Morgan, Charles. "Responsible AI: A Global Policy Framework." *ITechLaw* (November 18, 2019): 20. https://www.itechlaw.org/ResponsibleAI.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York, NYU Press, 2018). http://ebookcentral.proquest.com/lib/cornell/detail.action?docID=4834260.
Nourbakhsh, Illah Reza. *Robot Futures*. Cambridge, MA: MIT Press, 2013.
O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Broadway Books, 2016. http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1109940&site=ehost-live.
Pasquale, Frank. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge: Harvard University Press, 2015, 320. https://doi.org/10.4159/harvard.9780674736061.
Paszti, Laila. "Did the OPC Get It Right? A Proposed Approach for Regulating Artificial Intelligence." Canada | Global law firm | Norton Rose Fulbright. 2020. https://www.nortonrosefulbright.com/en-ca/knowledge/publications/33076187/did-the-opc-get-it-right-a-proposed-approach-for-regulating-artificial-intelligence.
Shah, Julie A., and Melissa Nobles. "Working Group on Social Implications and Responsibilities of Computing Final Report." Cambridge, MA: MIT Schwarzman College of Computing Task Force, 2019, 1–22. http://web.mit.edu/comptfreport/sirc.pdf.
Stork, David G. "The Best-Informed Dream: Hal and the Vision of 2001." *Hal's Legacy: 2001's Computer as Dream and Reality*, edited by David Stork, 7–15. Cambridge, MA: MIT Press, 1997.
Vanschoren, Joaquin. "Democratizing Machine Learning." OpenML. 2020. https://www.openml.org/.
Waraksa, Elizabeth A., ed. "Research Library Issues. RLI 299: Ethics of Artificial Intelligence." Association of Research Libraries (2019): 5.

Chapter 14

Machine Information Behaviour

Michael Ridley

Introduction

Most services and resources in academic libraries are grounded in an understanding of human information behaviour. Collections, systems, programs, and processes acknowledge and influence the ways in which people "need, seek, manage, give, and use information in different contexts."[1] Effectively, the library and the academy are in service to human information behaviour (HIB).

While the importance of HIB will remain, the proliferation of machine learning (ML) systems presents a new challenge to academic library services and resources. Increasingly, academic libraries need to consider the implications of machine information behaviour (MIB) and how those behaviours influence the services, resources, and programs they offer. Understanding MIB is a response to Bourg's challenge that algorithms be viewed as "a new kind of patron" necessitating a transformation in the manner in which the library responds.[2]

Algorithmic decision-making systems are ubiquitous, powerful, sometimes opaque, often invisible, and, most importantly, consequential in our everyday lives.[3] As these systems become more autonomous, even if in restricted domains, they will be utilized for recommendations and predictions regarding increasingly complex problems. However, "the danger is not so much in delegating cognitive tasks, but in distancing ourselves from—or in not knowing about—the nature and precise mechanisms of that delegation."[4] Understanding MIB will be essential to assuring veracity and engendering the trust necessary for delegation and use.

This chapter presents a preliminary conceptual model of machine information behaviour as a starting point upon which to build further elaborations and contextualizations. Using

Wilson's general theory of information behaviour as a scaffold, the model will incorporate the main functional components of these systems (i.e., computation, data, and algorithms) while also positioning them in the social, political, and economic environments.[5] Prominent in the model will be the three core elements present and active in any ML system: representation, evaluation, and optimization.[6]

Academic Libraries and Artificial Intelligence

The pioneering work in the 1980s and 1990s from LIS scholars and practitioners such as Linda C. Smith, Charles W. Bailey, Karen Spärck Jones, and F. W. Lancaster explored practical applications for AI in academic libraries, including the use of expert systems for reference service and information retrieval.[7] Following the hiatus of the "AI winter," brought about by the limitations of expert systems, renewed LIS interest in AI began in the 2010s and has grown substantially in recent years.[8] That said, the majority of this work has again focused more on practical applications of AI and less on its foundational implications.

Examples of work investigating the larger implications for LIS include search, discovery, reference and collections, and information literacy.[9] These and other critiques of AI in LIS have identified various failures and shortcomings related to bias, unfairness, discrimination, and accuracy. Often, these are linked to training data (or its preparation) and generically to the algorithms in question. However, as cognitive delegation to machine learning increases in many aspects of academic libraries and librarianship, an analysis and understanding of the complete contextual implementation of machine learning is required. The specific techniques and strategies of machine learning utilized at various stages of model training have a material downstream effect on information behaviour.

Applying an information behaviour (IB) lens to machine learning allows for a deeper understanding of the nature and consequences of this technology. Just as human information behaviour has shaped academic libraries, so too will machine information behaviour be a critical factor and have a profound impact.

Machine Behaviour and Machine Information Behaviour

Foundational to MIB is the concept of machine behaviour, "the scientific study of behaviour exhibited by intelligent machines [involving] a class of actors with particular behavioural patterns and ecology [requiring] the integrated study of algorithms and the social environments in which algorithms operate."[10] The authors advocate for the use of human behaviour research methods for research into machine behaviour. They caution, however, that "even if borrowing existing behavioural scientific methods can prove useful for the study of machines, machines may exhibit forms of intelligence and behaviour that are qualitatively different—even alien—from those seen in biological agents."[11]

A critique of Rahwan et al. suggests that the fields of cybernetics, science and technology studies (STS), sociology, and anthropology have for years undertaken similar approaches.[12] In the specific area of IB, however, this is not the case. The description of machine behaviour by Rahwan et al. provides a framework for the study of MIB in the context of HIB, allowing for behaviours that are both similar and different. Arising from this, machine information behaviour can be defined in the same terms as human information behaviour: systems or agents that "need, seek, manage, give, and use information in different contexts."

Conceptual Models

A conceptual model "provides a working strategy, a scheme" comprised of concepts, components, relationships, events, and changes.[13] Stafford notes, "The usefulness of a model lies in how it informs us about the potential relationships between features of the world."[14] Box famously observed that "all models are wrong but some are useful" emphasizing their role as always incomplete and emergent maps that attempt to define causality and provide a context for further research.[15] As a result, "models must be built by an interactive feedback process in which an initial parsimonious model may be modified."[16]

Any MIB model must consider knowledge representations (symbolic, statistical, and subsymbolic), learning methods (supervised, unsupervised, self-supervised, and reinforcement learning), specific algorithms, computational environments, and data sources for training and use.[17] It must also include the sociotechnical aspects of algorithmic systems that include the political, economic, and social implications of this technology.[18] The proposed MIB model is a starting point for an ongoing assessment through the application of further empirical studies.

Artificial Intelligence, Explainable AI, and MIB

Artificial intelligence (AI) is a broad term encompassing a variety of theories, strategies, and techniques to accomplish intelligent systems. Different approaches are used to represent knowledge, assess accuracy, and optimize results. The information behaviours of these approaches exhibit both similarities and differences. Selecting a particular AI method to accomplish a task dictates the resulting MIB.

Expert systems leverage human expertise codified into rules and logic statements.[19] These systems are "brittle" because of their limited domain knowledge and difficulties in knowledge base updating. However, their processes and outcomes are highly transparent and open to inspection. Neural networks and deep learning systems utilize big data, complex algorithms, and extensive computation to make predictions and recommendations based on probabilistic models.[20] These systems are opaque; they lack transparency and resist explanation. Recently, ML models have been critiqued for their lack of contextual awareness.[21] All AI systems either balance computational power and human intervention or preference one of them.

In a provocative blog post, Rich Sutton, the leading proponent of reinforcement learning, claimed that "the biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin …the only thing that matters in the long run is the leveraging of computation."[22] Sutton's argument refutes the role of human knowledge engineering in AI. Preferencing computation accepts that in MIB, "intelligence is not an information problem, it's a computational problem."[23] Allowing computers to maximize their specific strengths will generate processes and discoveries unmatched by humans and resulting in what Beatrice Fazi calls beneficial "alien thought."[24]

A bias in favour of computation, however, has contributed to the opacity of neural networks (*black boxes*).[25] If the information behaviours of these systems are largely opaque, what accountability measures are required to ensure veracity and to engender trust? The field of *explainable AI* (XAI) attempts to answer these questions through a variety of strategies, techniques, and process.[26] While XAI is largely the domain of computer science and engineering, there is a strong case for leadership from academic libraries and librarianship.[27] A model of machine information behavior is an XAI strategy because it provides an abstraction of a complex system with the goal of explaining concepts, relationships, and actions.

HIB and MIB

A number of general human information behaviour models have been proposed.[28] The model developed by Wilson over a number of years culminated in his 2016 "general theory" of human information behaviour and is used to illustrate the intersection of HIB and MIB.[29]

Wilson's HIB model can be redrawn to preserve the core concepts, recognize his separation of information processing and information use, reinforce the iterative nature of many of the components, and to put it in a format more emendable to overlaying the core functions of an AI system.[30]

Wilson's model has seven foundational concepts: person-in-context, information need, activating mechanisms, intervening variables, information seeking behaviours, information processing, and information use. Unique contributions of Wilson's model are the concepts of activating mechanisms and intervening variables. Activating mechanisms are enablers and contributing theories (e.g., stress/coping theory, risk/reward theory, social learning theory) that bridge the gap between context and information seeking and use. Intervening variables, initially called "barriers" and later expanded to include more general contextual variables (e.g., environment, role, demographic, psychological, and information source characteristics), identify influences that have a material impact on information behaviour, especially during the seeking and processing stages.[31]

These broad concepts and their interactions are sufficiently inclusive to account for the IB theories that focus on specific contexts and roles. The interactions among these concepts are non-linear. Activating mechanisms, intervening variables, and information seeking behaviours interact throughout an IB process or event. Similarly, information need, while an initiating event, is also a context that is refined throughout the seeking and use process.[32]

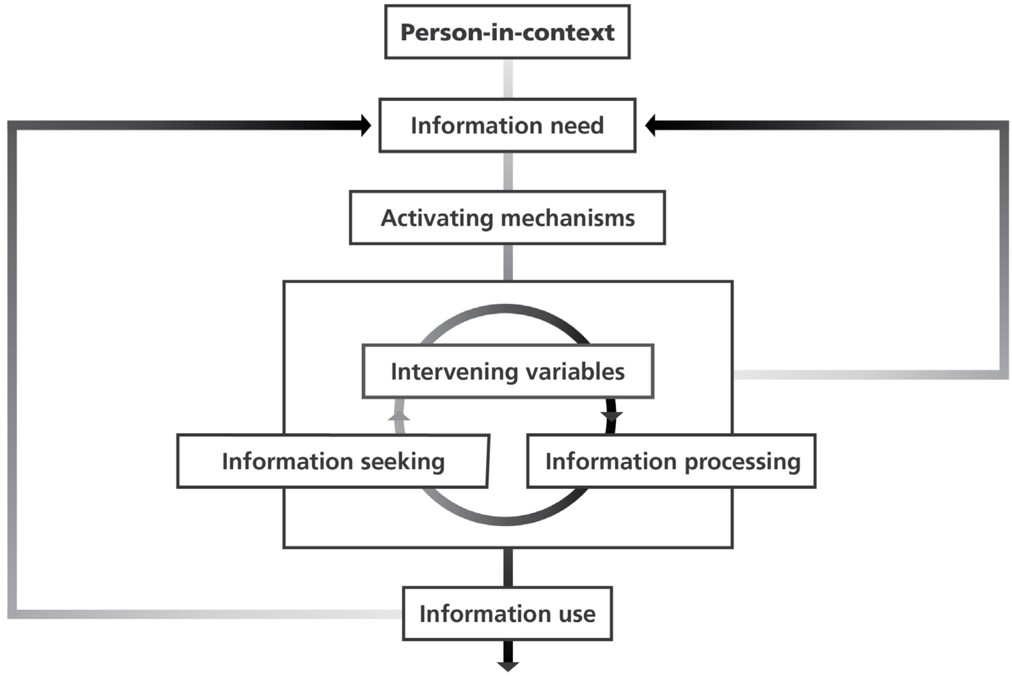


Figure 14.1
Wilson's General Theory of Information Behaviour (redrawn by the author).

There are three core functions common to all AI models: representation, evaluation, and optimization.[33]

Figure 14.2
Machine learning model.

Representation is how knowledge is expressed (e.g., rules, logic, vectors) as well as how the data is structured and understood. Evaluation is the scoring function of the model and how well the model fits the data. Optimization is the process that searches for the best model using specific testing and refinement techniques. The optimization and evaluation components iterate as model parameters and hyperparameters are adjusted and the result tested against the objective function (e.g., accuracy, similarity). While these processes are implemented differently according to the ontology that frames the intelligent machine (e.g., symbolic, statistical, subsymbolic), all are present and all influence MIB.[34]

By superimposing the core elements of machine learning on Wilson's general theory of information behaviour, the result is an illustrative and contextual interpretation of machine information behaviour.

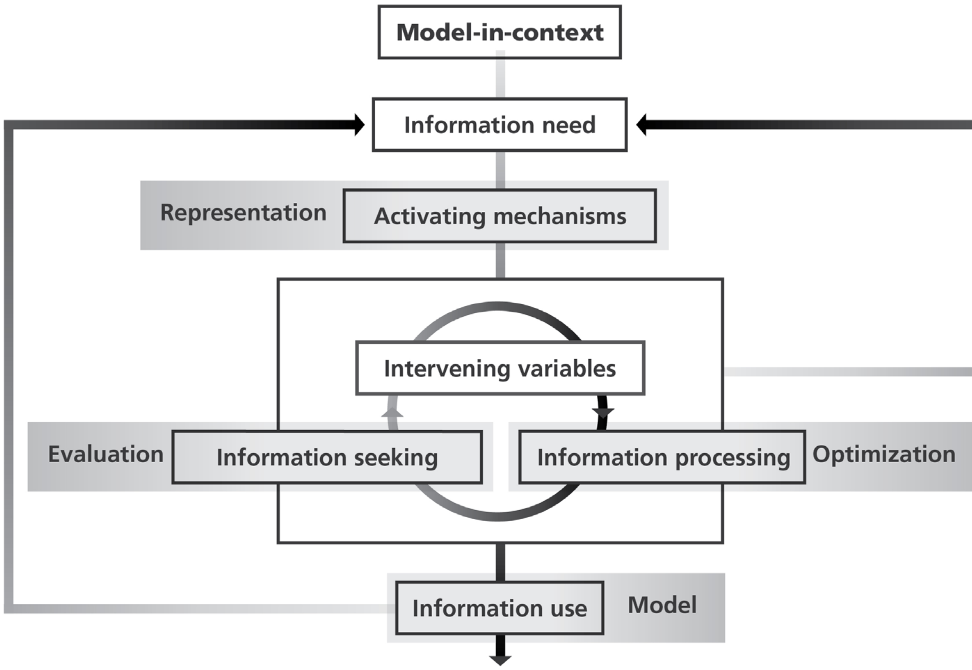


Figure 14.3
A preliminary machine information behaviour model.

All the components of this preliminary model can be elaborated to further define MIB. A brief examination of activating mechanisms, intervening variables, and information seeking and processing in the context of machine behaviour illustrates some of those characteristics.

Activating Mechanisms

An example of an activating mechanism in MIB is the ontology or paradigm at the core of the AI model. These consist of symbolists, connectionists, evolutionaries, Bayesians, and analogizers.[35] Each of these has a different concept of knowledge representation, learning methods, evaluation metrics, and optimization techniques. While not mutually exclusive, these ontologies prescribe or preference specific processes and representations that dictate subsequent actions and affect possible outcomes.

Another example is the process of data preparation, widely viewed as 80 percent of the effort in building a model, which cleanses and formats data in a manner consistent with the selected ontology (e.g., rules, vectors). This data preparation directly influences the subsequent evaluation metrics and methods as well as the optimization benchmarks and techniques.

Intervening Variables

As with the HIB model, intervening variables in MIB can come from a wide variety of sources, with each having different but material effects on information behaviour. For example, regulation and legislation may require systems to conform in specific ways. The global influence of the EU's General Data Protection Regulation (GDPR), with its notional "right to explanation," has driven widespread requirements for XAI.[36] Models (including assemblages into systems and agents) must be able to respond to demands for interpretability, transparency, and explainability.

Computation capacity is another important variable. The "combinatorial explosion" resulting from large information spaces can result in excessive computational demands.[37] Hence, computational availability and efficiency are significant variables that impact how long and to what depth the model can be trained.[38]

Information Seeking and Information Processing

These are iterative steps in both HIB and MIB. In HIB, these are approaches to resolving the information gap or need (e.g., active or passive search, passive attention, ongoing search) and to evaluate and synthesize that information for subsequent use. In MIB, these are the approaches to evaluation and optimization. Information seeking, intervening variables, and information processing come together to interrogate data, create hypotheses, and form (and test) interim models. Effective information processing is key to MIB, and a wide variety of strategies and techniques are employed. This aspect of MIB is an optimization process and is analogous to the stages in Dervin's sense-making model.[39]

AI-Authorship: An Example

In 2019, Springer Nature published *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research*.[40] The author is identified as "Beta Writer," an AI. The book production process, a collaboration between various machine learning processes and human editors, is fully documented in the introduction.[41] The book is an annotated bibliography of 151 key research publications in the field algorithmically selected, categorized, and summarized by "off-the-shelf" ML techniques and natural language processing (NLP) tools. It consists of four thematic chapters, each with an introduction, topic subsections with document summaries, conclusion, related works, and references. As an experiment in scholarly publishing, Springer Nature is fully transparent about the processes and decisions, successful and otherwise. The book is a useful example of MIB. Since it is not a fully autonomous machine learning process, the book is better viewed as a collaboration where the information behaviours reflect those of both humans and the machine learning algorithms.

The book production process can be seen as an iteration through the proposed MIB model (figure 14.3) while addressing specific tasks: preprocessing data, structure generation, text generation, and post-processing. Depending on the task and the iteration, core information behaviours can be identified and their implications recognized. Information need remains a human-directed behavior. However, aspects such as *activating mechanisms*, *information seeking*, *information processing*, and *intervening variables* illustrate MIB.

For example, various *activating mechanisms* include data preprocessing and setting similarity metrics for eventual topic clustering. Algorithmic processing of the approximately one thousand core documents for linguistic and semantic normalizations, the use of word embedding (a domain-specific issue in specialized areas, such as chemistry), and the production of the term-document matrix used to determine document similarity, all shaped the determination of chapters and their sections during the selection and clustering processes.

The iterations through *information seeking* and *information processing* illustrate the behaviour of the clustering and summarization algorithms. For example, the clustering algorithm and tuning for similarity sensitivity both impact document relatedness and confidence levels regarding inclusion in chapter sections. In generating chapter topics and then subsection topics within these, different clustering algorithms were tested (hierarchical clustering through tree structures and recursive non-hierarchical clustering). The latter was eventually used as the former resulted in lengthy processing times and uneven homogeneity among chapters. The structure of the book was algorithmically generated but, as with most ML systems, certain parameters were set and tuned by the editors (e.g., the target number of chapters and sections, the maximum number of documents per section, term frequency metrics, and the type of stemming and other normalizations used). The choice of another clustering algorithm, such as HDBSCAN, would have resulted in the autonomous determination of many of these parameters.

Document summarizations were drawn from the abstracts. A variety of techniques were used and critiqued by content experts: unsupervised extractive, supervised extractive summarization, extended abstracts (reformulated, compressed, and enriched), and a weighted combined ranking that utilized all three approaches. Ultimately, extended abstracts were used because of errors attributable to the other techniques and to the nature of the subject domain. While abstractive summarization is a preferred algorithmic approach, extractive summarization proved more reliable and readable.

Intervening variables can be identified by their presence and, in some cases, by their absence. Called a “minimalist implementation” by the book editors, this conservative approach resulted in the use of less complex algorithms and more moderate parameter settings to favour recall over precision and to enhance trustworthiness among the science community readership. A robust chemistry-specific ontology was not used, although examples such as the Springer Nature SciGraph would have been helpful. The availability and use of knowledge graphs (domain-specific as well as broader contextual mappings) are significant intervening variables in MIB. Human intervention in the algorithmic decisions was limited. Content experts moved only nine documents to different chapters and removed only eight from the final key research documents algorithmically selected.

The ML model used to generate the book has many hyperparameters set by humans and parameters learned by the algorithms. For example, these parameters directly affect the nature of the document selection and categorization as well as the manner in which the text summarization is constructed and presented. A future priority for the book is "to provide a user interface that allows a user to switch parameters on the fly and see and evaluate the modification obtained by this and thus optimize the machine-generated text according to personal preferences."[42] Such a dynamic reconstruction of the book would allow readers to impose their own tolerances for scope, precision and recall, and trustworthiness. This, in effect, would allow the user to modify the MIB of the machine learning model (i.e., the book).

Conclusion

Academic libraries, and the academy more generally, have both shaped and been shaped by human information behaviours. Artificial intelligence, through the significant advances of machine learning with neural networks and deep learning, has resulted in increasingly autonomous systems being used for complex predictions and recommendations. Bourg's "new patron" obligates academic libraries to understand machine information behaviour with the same attention previously applied to human information behaviour.

The challenges of ML are significant and well-documented. This is a technology with great promise and menacing peril. However, as de Mul and van den Berg observed, if cognitive delegation occurs, it must happen with a clear understanding of the nature, characteristics, and implications of the systems or agents we wish to use and trust.[43] The preliminary model of machine information behaviour presented here is merely a starting point. It is a means to focus attention on MIB and to position academic libraries and librarianship as a critical community for the exploration of this emerging field.

Endnotes

1. Karen E. Pettigrew, Raya Fidel, and Harry Bruce, "Conceptual Frameworks in Information Behavior," *Annual Review of Information Science and Technology (ARIST)* 35 (2001): 44.
2. Chris Bourg, "What Happens to Libraries and Librarians When Machines Can Read All the Books?," *Feral Librarian* (blog), March 17, 2017, https://chrisbourg.wordpress.com/2017/03/16/what-happens-to-libraries-and-librarians-when-machines-can-read-all-the-books/.
3. James Barrat, *Our Final Invention: Artificial Intelligence and the End of the Human Era* (New York: Thomas Dunne Books/St Martin's Press, 2013); Nicholas Carr, *The Glass Cage: Automation and Us* (New York: Norton, 2014); Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016); Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge, MA.: Harvard University Press, 2015).
4. Jos de Mul and Bibi van den Berg, "Remote Control: Human Autonomy in the Age of Computer-Mediated Agency," in *Law, Human Agency, and Autonomic Computing*, ed. Mireille Hildebrandt and Antoinette Rouvroy (Abingdon: Routledge, 2011), 59.
5. T. D. Wilson, "Information Behaviour: An Interdisciplinary Perspective," *Information Processing and Management* 33, no. 4 (1997): 551–72, https://doi.org/10.1016/S0306-4573(97)00028-9; T. D. Wilson, "A General Theory of Human Information Behaviour," *Information Research* 21, no. 4 (2016), http://www.informationr.net/ir/21-4/isic/isic1601.html.
6. Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (New York: Basic Books, 2015).
7. Linda C. Smith, "Artificial Intelligence and Information Retrieval," in *Annual Review of Information Science and Technology*, ed. Martha E. Williams (Amsterdam: Elsevier, 1987), 41–77; Charles

W. Bailey, "Intelligent Library Systems: Artificial Intelligence Technology and Library Automation Systems," ed. Joe A. Hewitt, *Advances in Library Automation and Networking* 4 (1991): 1–23; Karen Sparck Jones, "The Role of Artificial Intelligence in Information Retrieval," *Journal of the American Society for Information Science* 42, no. 8 (1991): 558–65; F. Wilfrid Lancaster and Linda C. Smith, eds., *Artificial Intelligence and Expert Systems: Will They Change the Library* (Urbana: Graduate School of Library Science, University of Illinois at Urbana-Champaign, 1992), http://hdl.handle.net/2027/inu.30000029957960.

8. Jason Griffey, ed., "Artificial Intelligence and Machine Learning in Libraries," *Library Technology Reports* 55, no. 1 (2019), https://doi.org/10.5860/ltr.55n1; Thomas Padilla, *Responsible Operations. Data Science, Machine Learning, and AI in Libraries* (Dublin, OH: OCLC Research, 2019), https://doi.org/10.25333/xk7z-9g97; Zehra Taskin and Umut Al, "Natural Language Processing Applications in Library and Information Science," *Online Information Review* 43, no. 4 (2019): 676–90, https://doi.org/10.1108/OIR-07-2018-0217; Andrew M. Cox, Stephen Pinfield, and Sophie Rutter, "The Intelligent Library: Thought Leaders' Views on the Likely Impact of Artificial Intelligence on Academic Libraries," *Library Hi Tech* (2018), https://doi.org/10.1108/LHT-08-2018-0105.
9. Safiya Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018); Matthew Reidsma, *Masked by Trust: Bias in Library Discovery* (Sacramento, CA: Litwin Books, 2019); Sofia Leung, Michelle Baildon, and Nicholas Albaugh, "Applying Concepts of Algorithmic Justice to Reference Instruction, and Collections Work" (Cambridge MA: MIT Libraries, 2019), https://dspace.mit.edu/handle/1721.1/122343; Alison J. Head, Barbara Fister, and Margy MacMillan, "Information Literacy in the Age of Algorithms: Student Experiences with News and Information, and the Need for Change" (Project Information Literacy, 2020), https://www.projectinfolit.org/uploads/2/7/5/4/27541717/algoreport.pdf.
10. Iyad Rahwan et al., "Machine Behaviour," *Nature* 568, no. 7753 (2019): 477–78, https://doi.org/10.1038/s41586-019-1138-y.
11. Rahwan et al., "Machine Behaviour," 483.
12. Emanuel Moss et al., "Machine Behaviour Is Old Wine in New Bottles," *Nature* 574 (2019): 176, https://doi.org/10.1038/d41586-019-03002-8.
13. Kalervo Järvelin and T. D. Wilson, "On Conceptual Models for Information Seeking and Retrieval Research," *Information Research* 9, no. 1 (2003), http://informationr.net/ir/9-1/paper163.html.
14. Tom Stafford, "What Use Are Computational Models of Cognitive Processes?," in *Connectionist Models of Behaviour and Cognition II: Proceedings of the Eleventh Neural Computation and Psychology Workshop, University of Oxford, UK, 16-18 July 2008*, ed. Julien Mayor, Nicolas Ruh, and Kim Plunkett, vol. 18, Progress in Neural Processing (New Jersey: World Scientific, 2009).
15. G. E. P. Box, "Robustness in the Strategy of Scientific Model Building" (paper, Madison: Mathematics Research Center, University of Wisconsin, 1979), 2.
16. Box, "Robustness in the Strategy," 30.
17. Francesco Corea, "AI Knowledge Map: How to Classify AI Technologies," Medium, August 29, 2018, https://medium.com/@Francesco_AI/ai-knowledge-map-how-to-classify-ai-technologies-6c073b969020; Domingos, *The Master Algorithm*.
18. Taina Bucher, *If…Then: Algorithmic Power and Politics* (New York: Oxford University Press, 2018).
19. Frederick Hayes-Roth, Donald Arthur Waterman, and Douglas B. Lenat, eds., *Building Expert Systems* (Reading, MA: Addison-Wesley, 1983).
20. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (Boston: MIT Press, 2016).
21. Gary Marcus and Ernest Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust* (New York: Pantheon, 2019); Hector J. Levesque, *Common Sense, the Turing Test, and the Quest for Real AI* (Cambridge, MA: MIT Press, 2017); Stuart Russell, *Human Compatible: Artificial Intelligence and the Problem of Control* (New York: Viking, 2019).
22. Richard S. Sutton, "The Bitter Lesson," *Incomplete Ideas* (blog), March 13, 2019, http://www.incompleteideas.net/IncIdeas/BitterLesson.html.
23. Sutton, "The Bitter Lesson."
24. M. Beatrice Fazi, "Can a Machine Think (Anything New)? Automation Beyond Simulation," *AI & Society* 34, no. 4 (2019): 813–24, https://doi.org/10.1007/s00146-018-0821-0.
25. Pasquale, *The Black Box Society*.

26. Tim Miller, "Explanation in Artificial Intelligence: Insights from the Social Sciences," *Artificial Intelligence* 267 (2019): 1–38, https://doi.org/10.1016/j.artint.2018.07.007.
27. Michael Ridley, "Explainable Artificial Intelligence," *Research Library Issues*, no. 299 (2019): 28–46, https://doi.org/10.29242/rli.299.3.
28. Natalya Godbold, "Beyond Information Seeking: Towards a General Model of Information Behaviour," *Information Research* 11, no. 4 (2006), http://www.informationr.net/ir/11-4/paper269.html; J. Liu, "Toward a Unified Model of Human Information Behavior: An Equilibrium Perspective," *Journal of Documentation* 73, no. 4 (2017): 666–88, https://doi.org/10.1108/JD-06-2016-0080; Diane Nahl, "Social–Biological Information Technology: An Integrated Conceptual Framework," *Journal of the American Society for Information Science and Technology* 58, no. 13 (2007): 2021–46, https://doi.org/10.1002/asi.20690; Bhuva Narayan, "Towards Developing an Integrated Model of Information Behaviour" (PhD dissertation, Brisbane, Queensland University of Technology, 2010), https://eprints.qut.edu.au/33252/1/Bhuvaneshwari_Lakshminarayanan_Thesis.pdf; Barbara Niedźwiedzka, "A Proposed General Model of Information Behaviour," *Information Research* 9, no. 1 (2003), http://informationr.net/ir/9-1/paper164.html; Andrew Robson and Lyn Robinson, "Building on Models of Information Behaviour: Linking Information Seeking and Communication," *Journal of Documentation* 69, no. 2 (2013): 169–93, https://doi.org/10.1108/00220411311300039; Andrew K. Shenton and Naomi V. Hay-Gibson, "Information Behaviour Meta-Models," *Library Review* 61, no. 2 (2012): 92–109, https://doi.org/10.1108/00242531211220735; Amanda Spink and Charles Cole, "Human Information Behavior: Integrating Diverse Approaches and Information Use," *Journal of the American Society for Information Science and Technology* 57, no. 1 (2006): 25–35, https://doi.org/10.1002/asi.20249; Pengyl Zhang and Dagobert Soergel, "Towards a Comprehensive Model of Cognitive Process and Mechanisms of Individual Sensemaking," *Journal of the American Society for Information Science* 65, no. 9 (2014): 1733–56, https://doi.org/10.1002/asi.23125.
29. T. D. Wilson, "On User Studies and Information Needs," *Journal of Documentation* 37, no. 1 (1981): 3–15, https://doi.org/10.1108/eb026702; Wilson, "Information Behaviour"; T. D. Wilson, "Models in Information Behavior Research," *Journal of Documentation* 55, no. 3 (1999): 249–70, https://doi.org/10.1108/EUM0000000007145; Wilson, "A General Theory of Human Information Behaviour."
30. Wilson, "A General Theory of Human Information Behaviour"; Wilson, "Information Behaviour."
31. Wilson, "Information Behaviour"; Wilson, "A General Theory of Human Information Behaviour."
32. Charles Cole, *Information Need: A Theory Connecting Information Search to Knowledge Formation*, ASIST Monograph Series (Medford, NJ: American Society for Information Science and Technology, 2012).
33. Domingos, *The Master Algorithm*.
34. Corea, "AI Knowledge Map."
35. Domingos, *The Master Algorithm*.
36. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016," European Union, 2016, http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679; Lilian Edwards and Michael Veale, "Slave to the Algorithm? Why a 'Right to Explanation' Is Probably Not the Remedy You Are Looking For" (SSRN, May 23, 2017), https://ssrn.com/abstract=2972855; Bryce Goodman and Seth Flaxman, "European Union Regulations on Algorithmic Decision Making and a 'Right to Explanation,'" *AI Magazine* 38, no. 3 (2017): 50–57, https://doi.org/10.1609/aimag.v38i3.2741; Margot E. Kaminski, "The Right to Explanation, Explained," *Berkeley Technology Law Journal* 34, no. 1 (2019): 189–218, https://doi.org/10.15779/Z38TD9N83H; Alejandro Barredo Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI," ArXiv:1910.10045 [Cs], October 22, 2019, http://arxiv.org/abs/1910.10045; DARPA, "Explainable Artificial Intelligence (XAI)" (Arlington, VA: DARPA, 2016), http://www.darpa.mil/attachments/DARPA-BAA-16-53.pdf; Andrés Páez, "The Pragmatic Turn in Explainable Artificial Intelligence (XAI)," *Minds and Machines*, 2019, https://doi.org/10.1007/s11023-019-09502-w.
37. Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge: Cambridge University Press, 2010).
38. Joel Hestness et al., "Deep Learning Scaling Is Predictable, Empirically," ArXiv:1712.00409 [Cs, Stat], December 1, 2017, http://arxiv.org/abs/1712.00409; Roy Schwartz et al., "Green AI," ArXiv:1907.10597 [Cs, Stat], August 13, 2019, http://arxiv.org/abs/1907.10597.

39. Brenda Dervin, *Sense-Making Methodology Reader: Selected Writings of Brenda Dervin* (Cresskill, NJ: Hampton Press, 2003).
40. Beta Writer, *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research* (Heidelberg: Springer Nature, 2019), https://link.springer.com/book/10.1007/978-3-030-16800-1.
41. Henning Schoenenberger, Christian Chiarcos, and Niko Schenk, "Preface," in *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research*, by Beta Writer (Heidelberg: Springer, 2019), v–xxiii, https://link.springer.com/content/pdf/10.1007%2F978-3-030-16800-1.pdf.
42. Schoenenberger, Chiarcos, and Schenk., "Preface," xxii.
43. de Mul and van den Berg, "Remote Control."

Bibliography

Arrieta, Alejandro Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García, et al. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI." ArXiv:1910.10045 [Cs], 2019. http://arxiv.org/abs/1910.10045.

Bailey, Charles W. "Intelligent Library Systems: Artificial Intelligence Technology and Library Automation Systems." Edited by Joe A. Hewitt, 1–23. *Advances in Library Automation and Networking* 4 (1991).

Barrat, James. *Our Final Invention: Artificial Intelligence and the End of the Human Era*. New York: Thomas Dunne Books/St Martin's Press, 2013.

Beta Writer. *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research*. Heidelberg: Springer Nature, 2019. https://link.springer.com/book/10.1007/978-3-030-16800-1.

Bourg, Chris. "What Happens to Libraries and Librarians When Machines Can Read All the Books?" *Feral Librarian* (blog), March 17, 2017. https://chrisbourg.wordpress.com/2017/03/16/what-happens-to-libraries-and-librarians-when-machines-can-read-all-the-books/.

Box, G. E. P. "Robustness in the Strategy of Scientific Model Building." Paper published by Madison: Mathematics Research Center, University of Wisconsin, 1979.

Bucher, Taina. *If…Then: Algorithmic Power and Politics*. New York: Oxford University Press, 2018.

Carr, Nicholas. *The Glass Cage: Automation and Us*. New York: Norton, 2014.

Cole, Charles. *Information Need: A Theory Connecting Information Search to Knowledge Formation*. ASIST Monograph Series. Medford, NJ: American Society for Information Science and Technology, 2012.

Corea, Francesco. "AI Knowledge Map: How to Classify AI Technologies." Medium. August 29, 2018. https://medium.com/@Francesco_AI/ai-knowledge-map-how-to-classify-ai-technologies-6c073b969020.

Cox, Andrew M., Stephen Pinfield, and Sophie Rutter. "The Intelligent Library: Thought Leaders' Views on the Likely Impact of Artificial Intelligence on Academic Libraries." *Library Hi Tech* (2018). https://doi.org/10.1108/LHT-08-2018-0105.

DARPA. "Explainable Artificial Intelligence (XAI)." Arlington, VA: DARPA, 2016. http://www.darpa.mil/attachments/DARPA-BAA-16-53.pdf.

de Mul, Jos, and Bibi van den Berg. "Remote Control: Human Autonomy in the Age of Computer-Mediated Agency." In *Law, Human Agency, and Autonomic Computing,* edited by Mirelle Hildebrandt and Antionette Rouvroy. Abingdon: Routledge, 2011.

Dervin, Brenda. *Sense-Making Methodology Reader: Selected Writings of Brenda Dervin*. Cresskill, NJ: Hampton Press, 2003.

Domingos, Pedro. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York: Basic Books, 2015.

Edwards, Lilian, and Michael Veale. "Slave to the Algorithm? Why a 'Right to Explanation' Is Probably Not the Remedy You Are Looking For." SSRN. May 23, 2017. https://ssrn.com/abstract=2972855.

European Union. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016," 2016. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679.

Fazi, M. Beatrice. "Can a Machine Think (Anything New)? Automation Beyond Simulation." *AI & Society* 34, no. 4 (2019): 813–24. https://doi.org/10.1007/s00146-018-0821-0.

Godbold, Natalya. "Beyond Information Seeking: Towards a General Model of Information Behaviour." *Information Research* 11, no. 4 (2006). http://www.informationr.net/ir/11-4/paper269.html.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Boston: MIT Press, 2016.

Goodman, Bryce, and Seth Flaxman. "European Union Regulations on Algorithmic Decision Making and a 'Right to Explanation.'" *AI Magazine* 38, no. 3 (2017): 50–57. https://doi.org/10.1609/aimag.v38i3.2741.

Griffey, Jason, ed. "Artificial Intelligence and Machine Learning in Libraries." *Library Technology Reports* 55, no. 1 (2019). https://doi.org/10.5860/ltr.55n1.

Hayes-Roth, Frederick, Donald Arthur Waterman, and Douglas B. Lenat, eds. *Building Expert Systems*. Reading, MA: Addison-Wesley, 1983.

Head, Alison J., Barbara Fister, and Margy MacMillan. "Information Literacy in the Age of Algorithms: Student Experiences with News and Information, and the Need for Change." Project Information Literacy, 2020. https://www.projectinfolit.org/uploads/2/7/5/4/27541717/algoreport.pdf.

Hestness, Joel, Sharan Narang, Newsha Ardalani, Gregory Diamos, Heewoo Jun, Hassan Kianinejad, Md Mostofa Ali Patwary, Yang Yang, and Yanqi Zhou. "Deep Learning Scaling Is Predictable, Empirically." ArXiv:1712.00409 [Cs, Stat]. December 1, 2017. http://arxiv.org/abs/1712.00409.

Järvelin, Kalervo, and T. D. Wilson. "On Conceptual Models for Information Seeking and Retrieval Research." *Information Research* 9, no. 1 (2003). http://informationr.net/ir/9-1/paper163.html.

Kaminski, Margot E. "The Right to Explanation, Explained." *Berkeley Technology Law Journal* 34, no. 1 (2019): 189–218. https://doi.org/10.15779/Z38TD9N83H.

Lancaster, F. Wilfrid, and Linda C. Smith, eds. *Artificial Intelligence and Expert Systems: Will They Change the Library*. Urbana: Graduate School of Library Science, University of Illinois at Urbana-Champaign, 1992. http://hdl.handle.net/2027/inu.30000029957960.

Leung, Sofia, Michelle Baildon, and Nicholas Albaugh. "Applying Concepts of Algorithmic Justice to Reference Instruction, and Collections Work." Cambridge MA: MIT Libraries, 2019. https://dspace.mit.edu/handle/1721.1/122343.

Levesque, Hector J. *Common Sense, the Turing Test, and the Quest for Real AI*. Cambridge, MA: MIT Press, 2017.

Liu, J. "Toward a Unified Model of Human Information Behavior: An Equilibrium Perspective." *Journal of Documentation* 73, no. 4 (2017): 666–88. https://doi.org/10.1108/JD-06-2016-0080.

Marcus, Gary, and Ernest Davis. *Rebooting AI: Building Artificial Intelligence We Can Trust*. New York: Pantheon, 2019.

Miller, Tim. "Explanation in Artificial Intelligence: Insights from the Social Sciences." *Artificial Intelligence* 267 (2019): 1–38. https://doi.org/10.1016/j.artint.2018.07.007.

Moss, Emanuel, Rumman Chowdhury, Bogdana Rakova, Sonya Schmer-Galunder, Reuben Binns, and Andrew Smart. "Machine Behaviour Is Old Wine in New Bottles." *Nature* 574 (2019): 176. https://doi.org/10.1038/d41586-019-03002-8.

Mul, Jos de, and Bibi van den Berg. "Remote Control: Human Autonomy in the Age of Computer-Mediated Agency." In *Law, Human Agency, and Autonomic Computing*, edited by Mireille Hildebrandt and Antoinette Rouvroy, 46–63. Abingdon: Routledge, 2011.

Nahl, Diane. "Social–Biological Information Technology: An Integrated Conceptual Framework." *Journal of the American Society for Information Science and Technology* 58, no. 13 (2007): 2021–46. https://doi.org/10.1002/asi.20690.

Narayan, Bhuva. "Towards Developing an Integrated Model of Information Behaviour." PhD dissertation, Queensland University of Technology, 2010. https://eprints.qut.edu.au/33252/1/Bhuvaneshwari_Lakshminarayanan_Thesis.pdf.

Niedźwiedzka, Barbara. "A Proposed General Model of Information Behaviour." *Information Research* 9, no. 1 (2003). http://informationr.net/ir/9-1/paper164.html.

Nilsson, Nils J. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press, 2010.

Noble, Safiya. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown, 2016.

Padilla, Thomas. *Responsible Operations. Data Science, Machine Learning, and AI in Libraries*. Dublin, OH: OCLC Research, 2019. https://doi.org/10.25333/xk7z-9g97.

Páez, Andrés. "The Pragmatic Turn in Explainable Artificial Intelligence (XAI)." *Minds and Machines* (2019). https://doi.org/10.1007/s11023-019-09502-w.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press, 2015.
Pettigrew, Karen E., Raya Fidel, and Harry Bruce. “Conceptual Frameworks in Information Behavior.” *Annual Review of Information Science and Technology (ARIST)* 35 (2001): 43–78.
Rahwan, Iyad, Manuel Cebrian, Nick Obradovich, Josh Bongard, Jean-François Bonnefon, Cynthia Breazeal, Jacob W. Crandall, et al. “Machine Behaviour.” *Nature* 568, no. 7753 (2019): 477–86. https://doi.org/10.1038/s41586-019-1138-y.
Reidsma, Matthew. *Masked by Trust: Bias in Library Discovery*. Sacramento, CA: Litwin Books, 2019.
Robson, Andrew, and Lyn Robinson. “Building on Models of Information Behaviour: Linking Information Seeking and Communication.” *Journal of Documentation* 69, no. 2 (2013): 169–93. https://doi.org/10.1108/00220411311300039.
Russell, Stuart. *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking, 2019.
Schoenenberger, Henning, Christian Chiarcos, and Niko Schenk. “Preface.” In *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research*, by Beta Writer, v–xxiii. Heidelberg: Springer, 2019. https://link.springer.com/content/pdf/10.1007%2F978-3-030-16800-1.pdf.
Schwartz, Roy, Jesse Dodge, Noah A. Smith, and Oren Etzioni. “Green AI.” ArXiv:1907.10597 [Cs, Stat]. August 13, 2019. http://arxiv.org/abs/1907.10597.
Shenton, Andrew K., and Naomi V. Hay-Gibson. “Information Behaviour Meta-Models.” *Library Review* 61, no. 2 (2012): 92–109. https://doi.org/10.1108/00242531211220735.
Smith, Linda C. “Artificial Intelligence and Information Retrieval.” In *Annual Review of Information Science and Technology*, edited by Martha E. Williams, 41–77. Amsterdam: Elsevier, 1987.
Sparck Jones, Karen. “The Role of Artificial Intelligence in Information Retrieval.” *Journal of the American Society for Information Science* 42, no. 8 (1991): 558–65.
Spink, Amanda, and Charles Cole. “Human Information Behavior: Integrating Diverse Approaches and Information Use.” *Journal of the American Society for Information Science and Technology* 57, no. 1 (2006): 25–35. https://doi.org/10.1002/asi.20249.
Stafford, Tom. “What Use Are Computational Models of Cognitive Processes?” In *Connectionist Models of Behaviour and Cognition II: Proceedings of the Eleventh Neural Computation and Psychology Workshop, University of Oxford, UK, 16-18 July 2008*, edited by Julien Mayor, Nicolas Ruh, and Kim Plunkett, Vol. 18. Progress in Neural Processing. New Jersey: World Scientific, 2009.
Sutton, Richard S. “The Bitter Lesson.” *Incomplete Ideas* (blog), March 13, 2019. http://www.incompleteideas.net/IncIdeas/BitterLesson.html.
Taskin, Zehra, and Umut Al. “Natural Language Processing Applications in Library and Information Science.” *Online Information Review* 43, no. 4 (2019): 676–90. https://doi.org/10.1108/OIR-07-2018-0217.
Wilson, T. D. “A General Theory of Human Information Behaviour.” *Information Research* 21, no. 4 (2016). http://www.informationr.net/ir/21-4/isic/isic1601.html.
———. “Information Behaviour: An Interdisciplinary Perspective.” *Information Processing and Management* 33, no. 4 (1997): 551–72. https://doi.org/10.1016/S0306-4573(97)00028-9.
———. “Models in Information Behavior Research.” *Journal of Documentation* 55, no. 3 (1999): 249–70. https://doi.org/10.1108/EUM0000000007145.
———. “On User Studies and Information Needs.” *Journal of Documentation* 37, no. 1 (1981): 3–15. https://doi.org/10.1108/eb026702.
Zhang, Pengyl, and Dagobert Soergel. “Towards a Comprehensive Model of Cognitive Process and Mechanisms of Individual Sensemaking.” *Journal of the American Society for Information Science* 65, no. 9 (2014): 1733–56. https://doi.org/10.1002/asi.23125.

Biographies

About the Editors

Sandy Hervieux is the virtual reference coordinator and the liaison librarian for political science, philosophy, and the School of Religious Studies at McGill University's Humanities and Social Sciences Library in Montreal, Quebec, Canada. Her research interests include reference services, information literacy, and the impact of artificial intelligence on user services.

Amanda Wheatley is the management, business, and entrepreneurship librarian at McGill University. She also serves as an expert for artificial intelligence and gamification in the library's Digital Scholarship Hub. Amanda's research is focused on the intersection of AI technologies with information seeking behaviours.

About the Authors

Bryan Abendschein (PhD, University of Illinois at Urbana-Champaign) is an assistant professor in the School of Communication at Western Michigan University. Bryan researches the intersection of interpersonal communication and technology usage, specifically during times of stress and change. He is also a member of the Communication and Social Robotics Labs.

Mark Bell, senior digital researcher, works in the Research and Academic Engagement team at The National Archives, UK. In his time there, he has worked on probabilistic record linkage, blockchain, and handwritten text recognition. He is particularly interested in applications of AI to archival collections and representing uncertainty.

Lynne Bowker is full professor at the University of Ottawa in Canada where she holds a cross-appointment between the School of Information Studies and the School of Translation and Interpretation. She was the 2019 researcher-in-residence at Concordia

University Library. Her teaching and research interests include machine translation and digital literacy.

Jairo Buitrago Ciro is a doctoral student in the PhD in digital transformation and innovation at the University of Ottawa. He holds a master of information studies from the University of Ottawa.

Catherine Nicole Coleman is digital research architect for the Stanford University Libraries and research director for Humanities+Design, a research lab at the Center for Spatial and Textual Analysis. Nicole works at the intersection of the digital library and digital scholarship as a lead architect in the design and development of practical research services.

Harrison Dekker is currently an associate professor and data services librarian at the University of Rhode Island where he directs the AI Lab and is involved in data science education. Prior to URI, Harrison enjoyed a long career at the UC Berkeley Libraries, where he developed the Library Data Lab.

Lise Doucette has been a librarian for fifteen years, working in the areas of assessment, data and statistics, research and instruction, and government information at Ontario universities.

Dr. Autumn P. Edwards (2006, Ohio University) is a professor in the School of Communication at Western Michigan University. Her scholarly interests include interpersonal communication and communication theory, in both face-to-face and computer-mediated contexts. Her research focuses on the influence of individuals' personal theories of communication on aspects of community life, relational health, and personal wellbeing. Autumn co-directs the Communication and Social Robotics Labs (www.combotlab.org).

Chad Edwards, PhD (2003, University of Kansas) is a professor of communication in the School of Communication at Western Michigan University. Edwards' research interests include human-machine communication, human-robot interaction, artificial intelligence, and instructional communication. Chad co-directs the Communication and Social Robotics Labs (www.combotlab.org).

Claudia Engel is academic technology specialist and lecturer in anthropology at Stanford University. She is a member of the Center for Interdisciplinary Digital Research at Stanford Libraries and holds a PhD in anthropology.

Juliana Espinosa is an assistant professor and user engagement librarian at the Western Michigan University Libraries. She earned her MLIS at San Jose State University and BA in anthropology at the University of California, Santa Barbara. She has presented on Myers-Briggs Type Indicator personality types of librarians, high-impact practices, and campus partnerships.

Angelica Ferria, curator at University of Rhode Island, University Libraries is a strategist and innovator focused on creating spaces and services that frame access to advanced technologies as a necessary component in connecting with, navigating, and utilizing increasingly complex information systems.

Charlie Harper is a data engineer at Flywheel Digital, where he works on big data problems as part of a leading tech organization. Previously, he spearheaded interdisciplinary digital projects at Case Western Reserve University's Freedman Center as the digital scholarship specialist, and guided data-driven research and technical approaches to cultural collections as a senior archaeologist at Florida's Bureau of Archaeological Research. His current research projects include the geospatial analysis of air and ground medical transportation and health outcomes, and studies of prehistoric architecture in Greece, Florida, and Ohio. Charlie holds a PhD in classical archaeology from Florida State University.

Maria Kalsatos is a master's student in the MA in Translation Studies program at Concordia University. She holds a BA in French language and literature from the University of Toronto.

Christina S. Y. Kim is liaison librarian at the Centre for Research & Innovation Support, a researcher-centric resource hub at the University of Toronto. Christina has held a variety of roles, including education and liaison librarian, communications lead, entrepreneurship librarian, and a cross-appointment as senior manager at MaRS Discovery District.

Ali Krzton is the research data management librarian for Auburn University Libraries, supporting data management and services across the research lifecycle. Ali has a BA from Dartmouth College, an MA from Texas A&M University, and a MLIS from Kent State University. She is interested in advancing the research data ecosystem.

Anne Kumer is an electronic resources metadata librarian at Case Western Reserve University. She collaborates with library departments to design and implement metadata schema to enhance search and discovery of digital collections. She holds an MLIS from Simmons College and a BA in English from the University of California, Berkeley.

Mike Maddalena is the director of data science and business intelligence at the Raymond A. Mason School of Business at William & Mary, where he leads the Business Intelligence team.

Indrani Mandal is a lecturer for the Department of Computer and Statistics at the University of Rhode island. She is the education coordinator for the URI AI Lab. The URI AI Lab is only a year old, but the lab has already been in the headlines for its undergraduate research and outreach programs. Indrani's interests are machine learning, data science, and AI education for K-12 students.

Evan Meszaros is a research services librarian at Case Western Reserve University. He provides reference consultations, information literacy instruction, and collection support for the science departments to which he is a liaison as well as for the broader university community. He holds an MS in biology from Case Western Reserve University and an MA in library and information studies from the University of Wisconsin-Madison.

Anna Milholland is the business librarian at the Raymond A. Mason School of Business at William & Mary, where she provides vision and leadership for the McLeod Business Library. Anna's research interests include service design, collection development, research data management, and digital scholarship.

Kim Nayyer is the Edward Cornell Law Librarian, associate dean for library services, and professor of the practice at Cornell University. Her research and teaching interests include critical information and data literacy, issues in data and machine learning, comparative copyright and knowledge protection, and diversity, inclusion, and decolonization issues. ORCID identifier 0000-0001-6733-4417.

Michael Ridley is librarian emeritus at the University of Guelph (Ontario, Canada) where he was previously the chief information officer (CIO) and chief librarian. He is currently a PhD candidate at the Faculty of Information and Media Studies, Western University (London, Ontario). Email: mridley@uoguelph.ca Twitter: @mridley

Marcelo Rodríguez works as a reference librarian in New York, NY. His research and writing interests include international relations and foreign and comparative law, particularly in Latin America and the Caribbean. He is an active member of the American Association of Law Libraries and the Canadian Association of Law Libraries.

Paulina Rousseau is a collections librarian, and liaison for arts, culture, and media at the University of Toronto Scarborough Campus. In 2014, she founded the Digital Pedagogy Institute, a multi-institutional annual conference focuses on the use of technologies to enhance and transform undergraduate and graduate teaching.

Amy Ruskin is a master's student in the School of Information Studies at McGill University. She holds a BA in computer science from Pomona College.

Jonathan Scherger is an assistant professor at the Western Michigan University Libraries. He received his MLIS from Wayne State University and holds a BA from the University of Michigan. His research interests include user experience, management in libraries, and assessment. He has presented on approaches to library programming, technology, and reference referrals.

Michael Serafin is a library technologies and liaison librarian at the University of Toronto Mississauga. As he is responsible for library technology and the liaison librarian for computer science and the Institute for Communication, Culture, Information and Technology, he is keen to explore new technology that supports students and faculty.

Jae Duk Seo is a multidisciplinary researcher and currently a second-year master's student in computer science. Jae is an AI project assistant in the Ryerson University Library Collaboratory, where he provides consultation and project support to Ryerson researchers and graduate students for AI-related initiatives. He has also co-led a popular program AI Journal Club in the Collaboratory. He is a researcher in the discipline of machine learning and computer vision, specifically on real-time keypoint detection and connection to information theory. He is a prolific content creator of open access educational materials focused on AI and machine learning.

Michelle Spence is an engineering and computer science librarian at the University of Toronto. She holds a HBSc and a MISt, both from the University of Toronto. She has held positions in academic and public libraries as well as a corporate setting.

Shelby Stuart is an early-career electronic resources librarian at Cleveland Clinic. Her experience ranges from public-facing to technical service roles. She works to facilitate and maintain access to an extensive collection of electronic resources. She holds an MSI from the University of Michigan and a BA in Sociology from Kalamazoo College.

Leontien Talboom is a collaborative PhD student at The National Archives, UK and University College London. Her research focuses on the constraints that digital preservation practitioners are facing when making born-digital material accessible.

Melissa Terras is the professor of digital cultural heritage in the University of Edinburgh's College of Arts, Humanities, and Social Sciences. She previously directed UCL Centre for Digital Humanities. Terras is a fellow of the Alan Turing Institute and trustee of the National Library of Scotland. You can generally find her on twitter @melissaterras.

Hilary Thorsen is resource sharing librarian in the Access Services Department at Stanford Libraries. Previously she worked in the Metadata Department and holds an MS in library and information science.

Carey Toane is entrepreneurship librarian at the Gerstein Science Information Centre at the University of Toronto, where she supports current and future U of T startups in their ventures. A former journalist, copywriter, and business librarian, her current research focuses on the information habits and needs of academic entrepreneurs.

Aaron Trehub is the assistant dean for technology and research support at the Auburn University Libraries. He is responsible for overseeing all aspects of library technology at Auburn, including the integrated library and discovery system, software development, digital scholarship, digital collections, digital preservation, and the Innovation and Research Commons (I&RC).

Aaron Tucker is a sessional instructor at Ryerson University, where he is a member of the Ryerson University Library Collaboratory, and a research fellow at the Ryerson Centre for Digital Humanities. Currently, he is a PhD student in the Cinema and Media Studies department at York University as an Elia Scholar, a doctoral scholar in the VISTA Lab, and a 2020 SSHRC Joseph-Armand Bombardier recipient. His graduate work studies the cinema of facial recognition software, machine learning, and artificial intelligence.

Patricia Vander Meer is a professor and outreach and instruction librarian at Western Michigan University, Kalamazoo, MI. Her publications have addressed online and large group library instruction, faculty attitudes toward instruction modes and the ACRL *Framework*, usability, designing instructional facilities, library website development, and utilizing robots in the academic library.

Fangmin Wang is an academic librarian and head of library information technology services (LITS) at Ryerson University. He oversees the library's digital technology strategy and operations. He has been instrumental in creating the Ryerson University Library Collaboratory, an interdisciplinary research incubation space that facilitates interdisciplinary collaboration and knowledge-sharing about the next generation of technology such as artificial intelligence. He was recognized as an innovator by *Library Journal*'s Movers and Shakers 2019 award program. He obtained his MLIS degree from McGill University.

Index

N

O

P

T

U

V

W

X

Y

Z