



This is a repository copy of *The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries.*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/137254/>

Version: Accepted Version

Article:

Cox, A.M. orcid.org/0000-0002-2587-245X, Pinfield, S. and Rutter, S. (2019) The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries. *Library Hi Tech*, 37 (3). pp. 418-435. ISSN 0737-8831

<https://doi.org/10.1108/LHT-08-2018-0105>

© 2018 Emerald Publishing Limited. This is an author produced version of a paper subsequently published in *Library Hi Tech*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

The intelligent library: Thought leaders' views on the likely impact of Artificial Intelligence on academic libraries

Abstract

Purpose (mandatory)

The last few years have seen a surge of interest in artificial intelligence (AI). The aim of this paper is to capture a snapshot of perceptions of the potential impact of AI on academic libraries and to reflect on its implications for library work.

Design/methodology/approach (mandatory)

Data for the study was interviews with 33 library directors, library commentators and experts in education and publishing.

Findings (mandatory)

Interviewees identified impacts of AI on search and resource discovery, on scholarly publishing, and on learning. Challenges included libraries being left outside the focus of development, ethical concerns, intelligibility of decisions and data quality. Some threat to jobs was perceived. A number of potential roles for academic libraries were identified such as data acquisition and curation, AI tool acquisition and infrastructure building, aiding user navigation and data literacy.

Originality/value (mandatory)

This is one of the first papers to examine current expectations around the impact of AI on academic libraries. The authors propose the paradigm of the intelligent library to capture the potential impact of AI for libraries.

Introduction

Following several years of intense activity around big data, there has been a surge of interest in AI. For example, in the UK, reports by the House of Commons Science and Technology Committee (2016) on AI and by the Royal Society (2017) specifically on Machine Learning, have been followed by the publication of findings of a House of Lords Select Committee on Artificial Intelligence (House of Lords, 2018). AI has come into public awareness through maturing consumer products that use voice recognition, such as Siri, and high profile innovations, such as smart cars (Tredinnick, 2017). Political interest in AI, motivated by its potential to raise productivity and stimulate economic growth, has been coupled with societal "AI anxiety" about the impact on jobs and social equality, and with a growing awareness of the risks to privacy (Johnson and Verdicchio, 2017). A Price Waterhouse Cooper report of 2017 found 54% of senior executives were planning to make major strategic investment in AI, but most thought their organisation currently lacked relevant skills (quoted Rao, 2017) pointing to another important issue, one echoed by Gartner (Andrews and Austin, 2018).

AI has a long history of development, but it seems to be on the cusp of a breakthrough in application. Some information sectors such as law are already starting to see a significant impact (Smith, 2016; Chen and Neary, 2017). The likely effects on academic libraries are unclear, however. In some senses, AI has already had effects here e.g. changes to search and discovery, experiments with chatbots and work supporting Text and Data Mining; though these are rarely understood as interconnected changes. That there will be further impacts on libraries seems inevitable. Fernandez

(2016:22), for instance, goes as far as to say, “For libraries the question is not so much what technology will be affected, but rather what technology, if any, will remain unaffected by AI”.

In a 2017 survey of librarians from across all sectors in the USA, Wood and Evans (2018) found that 56.3% of respondents thought supercomputers, like Watson, could transform librarianship. This still meant 44% thought it would have no or not much effect. Furthermore, most thought it would be 30 years before supercomputers would be in libraries. The effect would be seen in virtual services, discovery, referencing and cataloguing; other library functions would be less affected, respondents thought. Is that correct? And this leaves open the question of what might be the effects of other aspects of AI, beyond the model of Watson that Wood and Evans chose. Respondents saw the effect as mostly positive and not likely to involve the replacement of librarians or disintegration of the library. Is this optimism justified, when widely cited studies on the impact of automation more generally are more pessimistic? In their seminal study, Frey and Osborne (2017) estimate the probability of the replacement by computers of “library technicians” as 99%, “Library assistants, clerical” 95%, archivists 76% and librarians 65%.

In this context, the current paper seeks to capture a snapshot of views in 2017 on the potential impact of AI on academic libraries and to reflect on its implications for library work, based on interviews with 33 library directors, library thought leaders and experts from related areas.

Artificial Intelligence

Definition and scope

AI has long been an important area of research in computing. There have been previous spurts of development e.g. in the 1980s followed by “AI winters” (House of Lords, 2018). But towards the end of the current decade it seems that AI is “entering a crucial stage in its development and adoption” (House of Lords, 2018; 15). AI is not a unitary concept, however: it is usual to differentiate general or strong AI (aspiring to match the general intelligence of a human being) from narrow or weak AI where applications work on a particular problem space. It is in the latter where current development is happening.

Tredinnick (2017: 37) defines AI as “a cluster of technologies and approaches to computing focussed on the ability of computers to make flexible rational decisions in response to unpredictable environmental conditions.” Hare and Andrews (2017) define it as “systems that change behaviours without being explicitly programmed based on data collected, usage analysis and other observations.” It is a trend linking “process automation”, the “Internet of Things”, “Data processing”, “tangible robotics”, “conversational interaction” and “decision support,” they suggest. Smith (2016:221), writing about the impact in the legal sector, sees the “AI bucket” somewhat more narrowly as consisting of:

- Big data
- Analytics
- Machine learning
- Natural language processing
- Data visualisation
- Decision logic

Thus the hype around AI, builds on the hype around big data in the last few years, for it is the combination of masses of data with computing power that creates the potential for new levels of AI. Specific areas of development are machine learning, which refers to the ability of computer systems

to improve in an operation through processing data independent of explicit programming, and natural language processing (NLP), the ability of computers to accept inputs other than structured text e.g. voice recognition.

Library applications of AI

AI is already recognised as an important trend in the legal information sector, but is only just coming into focus in academic libraries. An important area of impact is likely to be in search/resource discovery. Fernandez (2016) identifies potential impacts in terms of analysing big corpuses of data, creating metadata, translation of search and integrating search across content. Completely new ways of interacting with information, e.g. location-based search will also be part of the picture, he suggests. The 2017 NMC library trend report sees the time to adoption horizon for AI as four to five years (Adams Becker et al., 2017). Yet in recognising AI as an umbrella term, some aspects of AI have already touched academic libraries: in addition to the impact on search and discovery, two stand out, namely chatbots and Text and Data Mining. Chatbots (also known as digital assistants, virtual agents or intelligent agents) are computer programmes that can simulate an intelligent conversation, through text, speech or potentially through an embodied representation. Chatbots have been developed to answer directional and other predictable enquiries, but ultimately could be developed to answer reference queries. Existing literature develops the argument that chatbots have advantages in terms of 24/7 availability, consistency and patience in answering queries. They could be experienced as less intimidating for users. Chatbots fit patterns of mobile communication among students and can build on understanding of user needs from existing chat services (Vincze, 2017). Inevitably, there are issues around system limitations, user acceptance, and in fact there has only been limited adoption by libraries.

TDM uses algorithms to analyse large bodies of content for patterns and information that it would be hard to discover for a human reader. As the scale of published information increases, researchers have a growing need to use tools to mine content from the literature and there is evidence of this happening in some sectors e.g. biomedicine and chemistry (McDonald and Kelly, 2012). The main way academic libraries have been touched by TDM is through involvement in trying to clarify the legal situation and to negotiate licences that allow mining, because the law / licensing for text mining is “restrictive, fragmented and uncertain” (Caspers et al., 2017: 135). Academic libraries seem to have a potential role in identifying and licensing content and software and perhaps training in use of tools (Anderson and Craiglow, 2017; Dyas-Correia and Alexopoulos, 2014). But few libraries seem to have yet developed such services.

Other AI-related developments, linked to the wider scholarly communication and educational contexts within which libraries operate, are yet to make a significant impact on libraries but some of them could be transformative in the long-term. Priem (2013) has pointed to the potentially disruptive changes in scholarly communication enabled by AI. He suggests that traditional journal publishing systems are likely to be replaced by “a set of decentralized, interoperable services that are built on a core infrastructure of open data and evolving standards” (Priem, 2013: 438). Neylon (2012) has characterised this as a “network-enabled literature”. In this environment, filtering of content, currently enabled through peer review of individual papers for particular journals, will, Priem (2013) suggests, be superseded by “powerful, online filters” which “distil communities’ impact judgements algorithmically, replacing the peer-review and journal systems”. Such systems would enable discovery of content to be achieved automatically through algorithmically-directed continually-learning harvesters which push findings to researchers in an automated way. If such disruptive impacts were felt on academic communities and publishers, the place of the academic library would inevitably be affected, though it is hard to discern precisely how.

AI is also set to have a profound impact on education, and this is once again likely to affect library roles, if indirectly. AI in Higher Education implies such trends as personalised learning, performance monitoring and learning analytics, intelligent tutors and the smart or intelligent campus. Building on the last decade of work around learning analytics, it is likely that AI can be used to support learning. For example, Underwood and Luckin (2011) review a range of existing AI applications that could assist learning processes through personalisation and flexibility, in a potentially inclusive way.

Issues raised by AI

AI encompasses some exciting areas of development with positive benefits, and is often construed as an inevitable change. However, there are a number of very significant issues with it that have raised public concern. Privacy is central to these. Commercial companies amassing huge amounts of user data ostensibly to personalise user experience, but also to target marketing, raise complex issues around manipulation, personal privacy and consent. This is especially the case because the companies concerned may be based in other legal jurisdictions.

Another important area of concern is bias: how far can AI be trusted to make fair decisions (AI report, 2017)? Given the complexity of the algorithms it becomes difficult to make the process of decisions intelligible. How are AI systems to be accountable and transparent if their operation cannot be understood? There is gathering evidence of the biased assumptions built into many algorithms, e.g. created through choice of training data. This may not merely be a teething problem; it can also be seen as related to structural issues in the AI industry, such as the preponderance of male employees, and the origins of funding for AI from state, including the military, and profit-driven commercial organisations.

Also AI is likely to be expensive, Luckin (2017) implies that cost is also a major barrier to AI in HE. If it comes to HE it may be primarily in the form of proprietary systems. If a commercial ethos drives development, this itself may be the most problematic aspect for Higher Education (Popenici and Kerr, 2017). All these considerations point to the need for the public to have greater data literacy, for wider discussion of the issues, and for improved transparency of systems and their design.

AI is likely to have an impact both on employment and social equality. There is a potentially positive story around the way that AI implies that librarians' knowledge of user needs, understanding of building collections, including licensing material, may put them in a strong position to play a role in creating AI infrastructure (Chen and Neary, 2017). This could mean the creation of new jobs. Librarians may be considered well placed to manage the process of introducing AI to avoid privacy pitfalls and help users to develop critical information literacy. If AI did impact work in the library sector, it could be to free up professional staff time from more repetitive tasks. Yet there is clearly also a risk of direct replacement of library staff by AI. Predictions already referred to by Frey and Osborne (2017) make grim reading for those in more routine roles.

Arlitsch and Newell (2017: 794) write that "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." Yet they acknowledge that the pattern of change remains unclear. Attitudes, understanding and expectations in the academic library sector are not well understood. Reference has already been made above to the snapshot of librarians' view provided Wood and Evans (2018). A survey of UK librarians' expectations of the future found that few respondents rated "AI and machine learning" as a highly significant trend, relative to a range of others (Pinfield et al. 2017). This paper analyses qualitative data to begin to build up a clearer picture of how the potential impacts of AI are perceived. More specifically, it answers three research questions:

1. What do library managers, library commentators and experts in education and publishing see as the implications of AI for academic libraries?
2. What issues do they perceive as arising from these applications?
3. What roles might academic libraries play in supporting and using AI?

Methodology

The data reported in this paper was collected in interviews with stakeholders both from within and beyond the academic library community. These participants could be broadly categorised into three groups: “Library Managers”, comprising directors of service and other senior academic library staff; “Library Commentators”, comprising academics, consultants and other experts in the field; and “Non-Library Participants”, a variety of thought leaders in higher education and technology-related organisations. Such categorisation was not always straightforward as participants carried out a wide range of roles and came from a variety of backgrounds, but the categories are used to give some context to the remarks reported. However, our analysis did not suggest any consistent patterns of stance by group. The interviewees consisted of a total of 33 participants: 23 from the UK, 10 international; 15 women, 18 men.

The interviews were conducted between May and July 2017 with each typically lasting an hour. Voluntary, informed consent was gained from participants, and the research approach gained ethical approval from the [anonymised institution] formal research ethics process. The interviews were wide-ranging and discussed the long-term future of academic libraries such as the effects of the digital shift, the continuing role of library space and perceptions of the future of the academic library. The interviews were recorded and fully transcribed. Thematic Analysis (Braun & Clarke 2006) was carried out on the interview transcripts, including a process of detailed coding, from which major themes in the data were identified. One strong theme was the potential impact of AI and it is the material from this theme that is presented in this paper. Six (out of 23) interviewees mentioned unprompted AI as one of the top 3 trends when they were asked to identify these at the beginning of the interview. All interviewees were also asked directly about the potential impact of “AI, machine learning and robotics” later in the interview; and in other questions about the future of academic libraries relevant material cropped up.

Implications of AI for academic libraries

AI and information discovery

Participants identified a range of impacts on search and resource discovery, some minor changes, others more fundamental. At a basic level, AI implied faster searches and also learning more about searcher behaviour. It could also improve search by identifying problems with metadata, reviewing how material was indexed and ensuring that it was discoverable:

“And machine learning would then say nobody can ever find it because it is catalogued in the wrong way or it is hidden in this particular area or, that it is not available electronically.” (Non-library participant)

Further, AI could also be used to produce metadata, perhaps in a less biased or certainly more efficient way than current manual methods. However, the whole approach to indexing material might also be changing:

“Pattern recognition. So for instance when you go into Google Photos and you can search Google photos for pictures of dogs and it will find every picture that

you have ever taken that has a dog in it. That is because Google's trained the neural network to recognise dog pictures and that's great for recognising patterns. So things that go beyond the traditional keyword search you know controlled vocabularies, metadata, and the sorts of things that we are used to in digital libraries." (Library commentator)

Recommendation systems would get better, and running in the background could even replace the need to search at all, because they would "anticipate your needs" (Library commentator). Amazon recommendation provided one model, but it needed to work better. Search results would come much more proactively to the user

"I think that will just grow massively in that research data will come to people's finger tips because of certain searches and algorithms in the background that are finding them stuff that they need." (Non-library participant)

Such recommendation could also be highly personalised:

"You could imagine providing a very distinct individual experience for each and every student and researcher based on AI." (Library commentator)

It might also lead to a diversification in how people were searching:

"So I think there is a lot more independence using the technology to support your own development for both sides students, staff and academics." (Library commentator)

In addition to the model of it working in the background to make recommendations based on user data, AI might at some point be able to conduct a type of reference interview.

"Certainly if AI gets to the point where it can interpret my questions in a way that is similar enough to the way that a librarian can interpret my questions. [...] The research interview isn't really happening as much, so AI potentially taking over from the research interview, as a way of asking the question "does your collection have something useful: something I want to go out and discover and do research with?" (Library commentator)

AI systems could then ultimately replace the current role of the library professional in conducting a 'live' reference interview, already seen as a declining activity.

Whilst participants often saw AI as potentially improving existing kinds of systems (supporting discovery) and enhancing current activities (such as metadata creation), there was also recognition of the possibility of these systems and activities actually being superseded by AI systems, transforming the ways in which people locate information of relevance to them.

The machine-readable collection

Mining text was recognised to be a key aspect of AI for the library sector. This implied that the future immediate user of the library collection would cease to be necessarily a human:

"So, I hope that as libraries start to understand the basics of machine learning that we would think through what the implications are for ensuring that our collections are accessible not just to human readers but to machines to read as well." (Library Manager)

Building libraries for an intelligent agent rather than a human reader had potentially profound implications for example around access:

“The assumption that the end goal is for human eyeballs to look at something will, perhaps not disappear, but reduce. Students and academics [are] using machines in the middle to bring back information in a condensed way. [...] That is going to have different implications on how you present the information, how you access it, how you licence it.” (Library commentator)

Readers might come to the collection with their own tools or it might be the library role to provide tools:

“Certainly, users of library collections be they of data or of literature are going to want to bring their own tools which will include machine learning and AI tools and apply them to the data, or to the digital content broadly. I think that smart libraries will start getting a lot more serious about how they can use machine learning and AI tools to improve discovery on behalf of their users.” (Library commentator)

One interviewee thought at this point the notion of a collection could change, indeed the nature of content could be transformed:

“They might completely upend that notion that the library attempts to licence or provide access to all of the publishers’ material. One can very easily imagine a situation where if I am procuring services for my research and student bodies and there is an ever growing amount of research out there that needs to be provided access to, but it is far more than any human could ever relate to. Do I instead focus in on licensing the very best machine learning, artificial intelligence services which are navigating that material [...] So I think they might change the nature of the collection to one where actually you are more interested in providing AI services for the student body rather than providing access to content. And actually does the nature of content as something distinct from a machine service stop existing?” (Library commentator)

Thus at some point a corpus of material sifted by AI tools for a user from the masses of published information, and existing only because of the configuration of that tool, could supersede the notion of a collection built by traditional mean.

Changes in research and scholarly communication

Interviewees thought that both research and scholarly communication were likely to change through AI, with potential implications for the library. Such changes in the long-term might disrupt the role of libraries, along with other stakeholders in the information environment, such as that of publishers.

There was a potentially profound impact on scholarly communication: AI had the potential to change peer review, the way journals work, even lead to the disappearance of publishers:

“AI is going to be super profound on scholarly communication, because it is going to replace quite of lot of what we do with peer review today and I am not just talking so much about the AI will be able to peer review the paper. We will see that for sure: especially checking validity of statistics, checking for coherence between figures and tables. [...] We are going to see a world in which a paper can

be published as a preprint, people can comment online, about that paper, and the AI will be able to look at the authority or the credibility of those people what the people have said, and from that pull an initial sort of peer review out right.” (Non-library participant)

The interviewee went on to argue that this form of automated peer review could then be customised to individuals:

“I think all of these factors will spell the end of journals as a way of communicating science because why should I subscribe to one particular journal or several particular journals that are trying to do their own independent filters when I could just subscribe to a filtering service that is exactly tailored to my interest as a researcher?” (Non-library participant)

This combination of quality filtering and discovery in highly-personalised forms delivered by AI could then have profound impacts on the research communication environment.

Changes in learning and teaching

There would also be effects on learning. There might be profound changes in course content e.g. in law and finance, where AI was seen to be having early impact on practice. Because curricula would have to change, so there would be an impact on the library, e.g. what data literacy might mean for a law student might be more like what a scientist needs.

Significantly library use data would feed into learning analytics:

“I think the data that libraries hold on user activity and the way they use services will contribute more widely to their kind of their learning journey in the future.” (Non-library participant)

AI could produce intelligent tools, such as ones to identify what is not well understood by learners and alert the teacher to the need to explain something better, and recommend useful resources:

“If you think of lecture capture, and 90% of students play the same 3 minutes of the lecture. And that is because the bit that the lecturer is talking about at that point nobody understands; it is really difficult. So machine learning could do two things it could say to the lecturer these 3 minutes are gone over and over again, so you need to change the way you do it. But it could also help by looking for resources. You may find, this useful. [...] Learning is customised to [the] individual” (Non-library participant)

Just as with personalisation of search and peer review, there could be profound personalisation of learning:

“IBM Watson being embedded into virtual learning environments to provide a kind of personalised, adapted learning experience.” (Non-library participant)

One interviewee imagined AI generating an entire curriculum for an individual:

“We are saying generate me a curriculum, generate me an exam paper tailored to this specific student’s requirements. Then it starts to get rather deep. The idea that a learner might end up pursuing what is really quite a unique pathway through the material because certain parts of the material have been amplified

because the lecturer feels that they need to do a bit more work in this area and a generative machine learning model is sat there essentially coming up with the challenges for them.” (Library commentator)

Once again the potential being identified seems to be around the automated personalisation of the experience, in this case of the learner, with mediation, in this case between the teacher and learner, being provided by AI systems.

It was recognised that through learning analytics students’ achievement would be monitored and measured. Equally there would be an impact on how library performance would be managed and measured.

“So how do we market our own value, in an environment where students are being consistently measured in terms of their progress? And if that is becoming the way that they measure our success how do we show our part of that as librarians?” (Library Manager)

Issues raised by AI

Participants recognised that AI also raised a host of challenging issues, both around how to apply it appropriately and through specific risks it posed. Most fundamentally, the library might not be a lead sector for many developments.

“The research in AI might be focusing on areas which are not actually able to take [advantage of] by librarians so for example production engineering.” (Library Manager)

Critically, the information sector could see major change but it would not necessarily be driven by libraries themselves, rather it would be other parties, such as commercial companies:

“But when you say the library, I don’t think of Leeds or Southampton or Imperial in 10 years having an AI based digital library. I think about Google having an AI based digital library.” (Non-library participant)

If this were the case the impact might be indirect, through changing expectations:

“People who use digital content are going to have increasingly greater expectations about what they can do with that content, in terms of discovery, analysis, annotation things like that.” (Library commentator)

“The capture of big data and the use of big data by massive services providers, people like Google, people like Facebook, the experience that they will deliver and how that raises expectations or alters expectations [...]. And related to that I think questions around privacy and what is done with one’s data.” (Library commentator)

Meeting such expectations could be challenging within library resource and capability.

The potential that AI was developed by big commercial companies was also linked to a fear that it might be a vehicle of marketization:

“I think that is going to be one of the biggest issues is marketisation of AI and machine learning. So you are losing that element of [...] what is good for academics and what is good for higher education and what is good for capitalism and what is good for these companies to make lots of money out of because the government is pushing that.” (Library commentator)

As AI is built on data, there would be a drive for connecting lots of sources of data about content and user behaviour, linked to the power that having such data would give its owner:

“So artificial intelligence and machine learning are only going to be able to informed through, they do learn but they also need huge amounts of data to learn. So if you have a one-off, they can’t help you.” (Library commentator)

In this context there were strong ethical concerns about how data to support AI was collected and used, for example in the context of learning analytics

“The thing that worries me about Google and Facebook etc is you are not the customer. Somebody else is the customer, another company is the customer you are the product, you are being mined for your personal information... and we see that I think already in the academic space.” (Library commentator)

Participants recognised the relevance of pre-existing debates such as that around filter bubbles and biases in data based services. Thus the issue of transparency and intelligibility of collection decisions arises:

“The idea that you might simply not know how a particular piece of content was found because a neural network recognised that it was in some ways similar to some other piece of content, that is quite a profound thing. You can’t really ask it and get an intelligible answer back all you can get is something that says hmm... it is kind of 85% similar to this one but you know take my word for it.” (Library commentator)

Trying to explain the results of search would become more of a challenge for information literacy. Ethical and privacy issues around how usage data was being used became critical:

“Machine learning is here. That is why we need to worry about the data and ethics stuff that is the library problem with AI, it is not just that it will come for your jobs, it is just that you are sleep walking into it by having a chat bot for your library. You are sleep walking into it by giving loads of money to Elsevier or Digital Science because Elsevier are a data broker.” (Library commentator)

If libraries did try to use data in the same way, would this be more accepted by users, especially if as some thought libraries might themselves be selling user data?

“I think there [are] also ethical issues whether libraries are ethical places and whether if you try and do more measuring or selling of data - which I think is happening a bit - if there [are] tensions there.” (Library Manager)

Applications such as learning analytics also implied strong issues around data quality and security:

“It will be a requirement to ensure that that data is correct and belongs to an individual so there may be more of a security issue and the ability to extract that data fairly quickly in real time.” (Non-library participant)

The issue of quality could also come up in the quality of research material:

“There are some examples of people publishing research papers that were created by these machine learning models. So we set the model to work and it created what anybody who is an expert in the field would regard as a load of nonsense and yet in some cases they have actually been published in peer reviewed journals. [...] How does the librarian specifically weed out this robo content let’s call it, this robotically generated stuff?” (Library commentator)

The changes around AI implied a need for new skills: new data literacies:

“But of course the big impact on skills and digital literacy, both for students and for academics and for librarians.” (Library Manager)

There was a need to develop skills around big data, to advise users and contribute to designing the infrastructure:

“Because library professional staff is much more now about metadata and managing databases and managing large volumes of data, and being able to support people to access that data and also to search and filter it. So I suspect the skills will get much more into the computer science end because, I mean 10 years from now, knowing what we know about things like IBM Watson, today, what is going to be happening with data and the way we access and evaluate it. You are right we probably will have librarians who almost have PhDs in computer science, because of the way that they will need to support particularly researchers in data mining and data sorting and data evaluation.” (Non-library participant)

“The modern librarian is a software engineer, who understands distribution mechanisms in the modern world. So it is not like the librarian disappears: that role of disseminating knowledge is more important than ever. I think what is shifting is what skills a librarian has to have to thrive in that modern world. It used to be that you need to understand the Dewey decimal system, and these technical skills like that, and maybe even some architectural concepts like the flow of people in the library and stuff like that. If you conceive the role of librarian as someone whose job it is to disseminate and curate academic information the skills there you need are software engineering, building tools online” (Non-library participant)

One participant articulated the fear that libraries’ risk averseness would hold them back from participating:

“So libraries are always reactive, rather than proactive when it comes to AI ((laughs)) or management of third spaces or, you know any of that stuff because they don’t want to be police and they don’t want to have radical ideas. And, they don’t want to push back on the institution about things like learning analytics or having sensors or collecting lots of data on people.” (Library commentator)

Interviewees recognised that there was a threat to jobs:

“So everything from ordering books to cataloguing books etc, answering queries. I think software is going to replace people. I think this whole chatbot piece, why do you need an individual, you can do lots more of these things in an automated self-service way? [...] There is lots of talk about the white collar job going under the automation hammer and I think that you can see that happening in libraries.”
(Non-library participant)

While some jobs were threatened, there was an optimistic aspect if it freed people to do more interesting things:

“When we talk about robots and AI people tend to focus on the dystopian aspects, and the flip side of that is opportunities for people who are doing things that are quite repetitive and often demeaning and menial to free up their time, free up their capabilities to do more worthwhile things.” (Library commentator)

Indeed, one interviewee felt that if a current service could be replaced by automation, that just showed it was being done wrongly already:

“So if you are delivering stuff in your library that could be delivered by a chat bot you are probably not actually doing the things that you should be doing with your humans.” (Library commentator)

Library roles in AI

The uses and issues discussed in the previous sections implied a range of potential roles for libraries, that interviewees also outlined. Although some participants spoke of changes that might bypass libraries and supersede the role of librarians, many highlighted many ways in which there was an element of continuity with existing practices. One area of work could be around collections. Libraries might have a role in procuring content (including data sources) with appropriate licences:

“You will have to be very good at understanding the implications of the licences, contracts that they sign with some of these providers. You will need to be very good at understanding what it is that we can do with the derived outputs from these materials as institutions will want to, potentially monetize some of those outputs.” (Library commentator)

Libraries might also themselves be data providers, digitising their own unique material to be mined:

“You know where we have special collections we have got very long runs of printed statistics, and government publications. [It is a] very important research tool that is often unique. We have gone to work with academics to digitise but also to convert them into useable data sets and enable big data research on them. So, I think in that case the library is becoming more of a platform.” (Library Manager)

There was also potentially a role in procuring the most appropriate AI tools:

“Focus in on procuring licencing the very best machine learning... artificial intelligence services which are navigating that material able to provide that, the material that you actually need.” (Library commentator)

The role of providing content and tools, might extend beyond procuring data mining tools to managing a whole data infrastructure:

“If we are going to have data flowing like water and electricity, through let's call it the intellectual plumbing of the university to meet the needs of the students, the researchers, the academics then the people who provide that. Then it needs to be people who understand that information, understand the data, understand what data and information you need to meet the needs of the different disciplines, the different subjects, and so on and who can basically design an information supply model for a particular institution.” (Library Manager)

Because libraries are knowledgeable about user needs and trusted they could be key intermediaries in building services based on AI. One important aspect of a potential role was working out how to apply AI appropriately to the sector:

“So in terms of artificial intelligence and machine learning where we are going with that, we need to skill up our staff now so they can understand what it can do, and how that is mapped to the needs of our institutions, and higher education generally.” (Library commentator)

There was also a need to support the navigation of the landscape of content and tools:

“Navigating skills are still needed. We are still the specialists in that. I think if we are able to change our mindset towards us being - I am not quite sure if it is the right word - being a gatekeeper. Maybe it is more about being a knowledge centre because if we can change the role where they see the library or the librarians as the ones that can tell them how to navigate within big data, and within all this machine learning.” (Library Manager)

The issues of data quality in research publications – what one interviewee had dubbed “robo-content” – suggested another type of role:

“And I think there is a real role for librarians also probably using some of this technology to help weed things out. Security guards almost? Not gatekeepers exactly. Arbiters of quality perhaps.” (Library commentator)

The same issue of data quality and security revolved as much about usage data as about research content:

“I think that there is a big thread of this that revolves around who students and scholars are willing to trust with their personal history and their privacy.” (Library commentator)

Teaching data literacy skills and teaching people to understand how to protect their privacy would be important, and perhaps integrated into the existing information literacy training role:

“So whether we end up teaching students instead of information management how to dig into data and data visualisation and data mining, whether we end up in that space is a possibility. ” (Library Manager)

“Shouldn't we be the bastions of information literacy and information privacy in an AI world?” (Library commentator)

There was also potentially a need for curating derived outputs of research like data mining:

“The curatorial role is absolutely essential... it is more important than ever as I say because there is just more content.” (Non-library participant)

Relatively little mention was made in the interviews of librarians undertaking data analysis themselves. But one quote captured the extent of the potential:

“What we did last year is hire 5 analysts: we have one for scholarly information, for bibliometrics and so on, and one for GIS data, one for text data, one for legal issues, and one for business processes. I think we need analysts in many areas and librarians.” (Library Manager)

This is one of the few interviewees who saw that analysis could be part of the library role. His remark connects the need to analyse library problems e.g. legal issues or to analyse scholarly output through bibliometrics and to analyse text in text mining. It might also be needed for analysing library data and learning analytics. This asks a bigger question of where data analysis fits into the professional role.

Discussion

Around 2016 there seems to have been a tipping point in AI linked to a coming together of developments around machine learning, natural language processing and visualisation (Smith, 2016), along with increasing recognition of its significance in government and the media. The umbrella term AI also seems to capture a range of possible significant trends impacting academic libraries such as in search/recommendation and personalisation, TDM and the collection as data, learning/library analytics. JISC’s characterisation of the effects of AI in the context of the intelligent campus as “wide and deep” is apt (Clay, 2018). Further it could link to the impact of other data related work such as RDM and analytics work in bibliometrics/altmetrics. Some elements of the “AI bucket” like robotics, do not appear to be perceived as immediately relevant. Chatbots probably are relevant, though they were not mentioned very much in these interviews. Rather, most comment in the interviews focussed on the impact on the user experience in resource discovery. There was relatively little reflection on the way that both library management decision making and evaluation of libraries could be based more on data. The Internet of Things was also not really mentioned. Nor the intelligent campus, though this concept relates strongly to library provision of learning space, which at other points of the interviews was thought to be a key library role today and in the future. Thus the picture from the interview data suggests that only part of the potential impact of AI is in focus. Only two or three interviewees conceptualised AI as a truly transformative change. Where this was seen, however, it was recognised that it challenged current ideas of the library in fundamental ways. Others conceptualised AI as a series of incremental changes, often enhancing existing services or ways of working, but an analysis of the cumulative impact of such changes could lead to them being seen as transformational.

AI is not new: some is already here, indeed impacts are recognisable in some long familiar trends or aspirations (automation, self-service, improved search, customisation, recommendation). This was apparent, but there did also seem to be some realisation of potentially significant change coming.

In a paradigmatic form, in terms of their impact on libraries, these changes constitute what one might call the “Intelligent library” – a term first used by James Clay (2018) in the context of the intelligent campus concept. There are recognised to be some major challenges in such changes, particularly around commodification, user privacy, and libraries ability to respond. There was a

threat to jobs, though this did not seem to be a paramount concern. Certainly there will be changing expectations of libraries, through users' experiences of AI and other technologies in other contexts.

It emerged that a number of interrelated roles for libraries could arise from AI (summarised in Table 1): around providing and acquiring content, procuring tools to explore content and data, and stewarding derived outputs; as well as monitoring quality of material in the collection. There would be a need to support users to navigate a new information landscape, evaluate content and manage their digital privacy. Libraries might have a place in helping to design the whole infrastructure based on their understanding of user needs, and to protect users' privacy and interests with their data.

Library roles in AI	Competencies that need to be extended	Alternative providers of service/ function
Procuring content for AI to work from (including both licensing and through open access)	Procurement and licensing of e-content	Publishers and other new intermediaries
Providing content	Digitisation, metadata provision	Publishers and other new intermediaries
Data quality control	Collection management	
Procuring AI tools	Procurement and licensing of software and services	IT departments, academic departments
Data curation (e.g. of derived data)	Collection management, digital preservation	Publishers and other new intermediaries
Designing data infrastructure to enable AI	Design of information discovery infrastructure	IT departments
Explaining how to navigate the new information environment	Understanding of the scholarly publishing landscape, including data creation processes	
Teaching critical data literacy: understanding how to evaluate AI tools and their results, and also protect one's own privacy	Information literacy	IT departments
Designing AI tools	N/A – outside normal library professional work	Academic departments, Publishers
Data analysis and writing algorithms	N/A	IT departments, academic departments

Table 1 Potential library roles in AI

However, it could be argued that other units on campus such as IT services might be in a stronger natural position to take on some of these roles: e.g. procuring or designing AI tools or designing the data infrastructure. There might well also be new types of competition from commercial suppliers to mediate access, integrate content and support use. There are also questions about whether libraries have the capacity to respond, because the changes are so fundamental that this may be impossible or because resources simply do not exist. It seems likely that some of the roles that lie close to existing activities such as collection and information literacy will be privileged in how libraries

choose to respond. It seems far less likely that libraries will take up roles around data analysis, than around collection or information literacy.

One challenging issue arising from the findings was that relating to the level at which activity takes place and the associated issue of commercialisation. The corollary of many of the observations of the participants is that activity in designing and delivering new AI-enabled services would naturally take place at a level that was supra-institutional: sometimes at national levels, but often at the network level. The capacity of the library community to coordinate and resource activity at this level may be constrained. Because of this there is a real possibility that many of the most innovative developments might be undertaken by commercial organisations, which will inevitably attempt to monetize products and services. Commercial providers may dominate the landscape, with libraries having limited bargaining power or even losing their role entirely.

Interviews reflected a lack of consensus about the meaning and significance of AI and considerable uncertainty about its potential impact. The picture above is pieced together from the remarks of all participants; few seemed to have a very clear idea of the whole picture. At least one interviewee (a library director) admitted candidly to not understanding what AI was. Perhaps part of the explanation for this lack of clarity lies in the complex and long-term nature of the change that lies behind “AI”. It might also reflect the fact the AI field is rapidly changing. AI, itself a “bucket” of developments, is clearly tied up with a whole set of other changes (technological, social, economic etc) which make predicting its impact particularly difficult. Thus, the interview question that was posed (“How will AI, machine learning and robotics impact on libraries in the long term?”) frames it as an issue for the future. Yet, as we have seen, the timeframe within which AI was thought about by different participants varied significantly. One (thinking of it principally as data mining) saw it as something that had been well established for two decades: the burning issue was stewardship of derived data. Another felt an urgent sense that AI was happening now, that libraries were “sleep walking into it” through adopting chatbots or failing to realise how some of their corporate partners were monetizing data about use and users. Another thought it was “under our noses” but not recognised as the key trend it was. Often it was talked about as the future, but often as an extension of things that are already here or having impact over a number of years, such as experiences of using Amazon, Google or Facebook (e.g. for personalisation) or even CCTV (as a previous model of surveillance). Other comments were much more speculative about potential futures. Perhaps not surprisingly only three interviewees thought strongly in terms of AI as disruptive change, e.g. where it implied the end of the collection as previously understood. This variation reflects a reality of futures in a changing time: impacts will be patchy across the sector and with different functions affected quite differently at different rates of change (Wood and Evans, 2018). There are effects directly on the library and effects via changes in user behaviour and institutional contexts. This is the reality of experiencing a nexus of change: its impact is broad and deep, its scope hard to visualise or articulate. This uncertainty reflected the evidence from the literature review: that we are in the midst of a long-running transformative change, stretching back to the beginning of computing, affecting the immediate present and with ramifications into the longer term future.

Conclusion

This paper is one of the first to analyse a substantial body of empirical data about perceptions of how the latest wave of AI is likely to affect academic libraries, the services they deliver and the territories they occupy. The analysis has combined the views of a number of participants, with varying views, to construct an overall picture. Even so, it is clear from the literature that there is a wider picture still, which is only beginning to emerge. With the growing prominence of AI in current

governmental policies and increasing scrutiny of the strategies of the search / social media giants, it is highly likely there will in the immediate term continue to be intensive debate and activity around AI in many countries. This is likely to generate greater discussion and development in a wide range of communities, including academic libraries. The legal sector has already felt some of these effects, and there may be something to learn from the experience of law libraries. Some changes related to AI have already happened in the academic sector, some have even become taken for granted. Others can only be guessed at. Yet bringing our data together with the literature, it emerges that AI should be seen as defining a nexus of change (Pinfield et al., 2017) that has “wide and deep” ramifications (JISC reference) in terms of:

1. What a library is, what a collection is and how to search for material. The library may increasingly be seen as data, accessed through AI, the scope of the collection as framed by the AI;
2. How established services are delivered, for example by chatbots and other intelligent agents;
3. What users expect of libraries: through expectations learned in other areas;
4. What information literacy is: the ability to navigate a new space of AI tools and data, and data literacies, including critical awareness of how to protect one’s own privacy;
5. Who users are: some users will be AI tools; human access to content will be remediated through content being summarised and partially analysed for them by machines;
6. What libraries know about users and so how the library is managed: because of management decisions based on use data, combined with other learning and research analytics;
7. How the library works with other internal and external partners and competitors, especially IT services and new third-party commercial services;
8. How library services are evaluated: again through wider and deeper data;
9. What skills librarians need: be that for licensing, evaluation of data analysis and visualisation tools or using such tools themselves;
10. Whether the library community can operate effectively at different levels beyond the institution: in order to design and deliver services which will serve international communities of scholars and students;
11. Indeed, whether we need librarians (because of chatbots, automated metadata creation tools etc) or libraries (because of alternative intermediaries) at all, at least as currently conceived.

Such changes constitute the extent of the implications of the paradigm of what one could call the intelligent library. Such changes represent a significant challenge to the future position of the library and information profession(s). Clearly there are potential demands in terms of converting existing knowledge sets or developing new competencies, investment of effort, when there are many other demands on resources.

The demands of responding to such changes may reveal a significant skills gap in the sector. We know there is already strong demand in the economy for data scientists for data analysis and visualisation. Perhaps some librarians will be required to develop these skills, or at least awareness of different techniques and how they need to be supported. These demands are a challenge because they prompt librarians to learn more about IT and quantitative data analysis, including statistics. In a relatively low-paid sector we may be unlikely to attract people with stronger STM backgrounds to the profession, while those in the profession are typically from an arts background. More optimistically, we can say there will also be a need for librarians as data curators to take on new data

related skills: procurement and licensing, data management, quality control, curation and stewardship. Skills in these areas are stronger, and there has already been a focus on developing relevant skills driven by RDM.

Nevertheless, we can reasonably anticipate in the foreseeable future a wave of disillusion countering some of the current hype around AI (Tredinnick, 2017; Marcus and Davis, 2018), with many of the promised benefits not arriving immediately, and further rounds of AI development being required. Further public debate about data and privacy may also create negativity, or at least caution, in wider society about AI. And of course, we can also anticipate strong continuities for academic libraries, at least in the medium term. For example, the wider interviews in the current study recognise the continuing need for traditional forms of collections, including books, and the importance of physical library spaces (Pinfield et al, 2017). The intelligent library will not supersede the familiar academic library in the easily foreseeable future, rather it will exist as a new paradigm of development. It also seems probable that the impact of AI will be different in different sectors of academic libraries, e.g. teaching-led institutions as against the research led. The many trends changing academic libraries make it hard to focus on what should be given priority, given limited resources. Yet AI seems to be one prominent area that should be a focus of significant attention.

References

- Adams Becker, S., Cummins, M., Davis, A., Freeman, A., Giesinger Hall, C., Ananthanarayanan, V., Langley, K., and Wolfson, N. (2017). *NMC Horizon Report: 2017 Library Edition*. Austin, Texas: The New Media Consortium.
- Anderson, C. B., & Craiglow, H. A. (2017). Text mining in business libraries. *Journal of Business & Finance Librarianship*, 22(2), 149-165.
- Andrews, W. and Austin, T. *Artificial intelligence primer for 2018*. Gartner report.
- Arlitsch, K., & Newell, B. (2017). 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(7), 789-798.
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), pp.77–101.
- Campolo, A., Sanfilippo, M., Whittaker, M., Crawford, K. and Selbst, A. (2017). *AI now 2017 report*. https://ainowinstitute.org/AI_Now_2017_Report.pdf
- Caspers, M., Guibault, L., McNeice, K., Piperidis, S., Pouli, K., Eskevich, M. and Gavriliidou, M. (2017). *Baseline report of policies and barriers of TDM in Europe (extended version)*. FutureTDM project report. https://www.futuretdm.eu/wp-content/uploads/FutureTDM_D3.3-Baseline-Report-of-Policies-and-Barriers-of-TDM-in-Europe-1.pdf
- Chen, SX and Neary, MA. (2017). Artificial intelligence: Legal Research and law librarians. *AALL Spectrum* 21 (5) 16-20.
- Clay, J. (2018). The challenge of the intelligent library. Keynote at What does your eResources data really tell you? 27th February, CILIP. <https://intelligentcampus.jiscinvolve.org/wp/2018/03/02/sherif-keynote/>
- Dyas-Correia, S., & Alexopoulos, M. (2014). Text and data mining: Searching for buried treasures. *Serials Review*, 40(3), 210-216.

- Fernandez, P. (2016). "Through the looking glass: envisioning new library technologies" how artificial intelligence will impact libraries. *Library Hi Tech News*, 33(5), 5-8.
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: how susceptible are jobs to computerisation?. *Technological forecasting and social change*, 114, 254-280.
- House of Commons Science and Technology Committee. (2016). *Robotics and artificial intelligence*. <https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2015/robotics-and-artificial-intelligence-inquiry-15-16/>
- House of Lords Select Committee on Artificial Intelligence. (2018). *AI in the UK: ready, willing and able?* House of Lords. <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>
- Johnson, D. G., & Verdicchio, M. (2017). "AI Anxiety". *Journal of the Association for Information Science and Technology*, 68(9), 2267-2270.
- Luckin, R. (2016). Four ways that artificial intelligence can benefit universities: Higher education needs to embrace the positives of AI, not just look at the negatives. *Times Higher Education*, August.
- McDonald, D. and Kelly, U. (2012). *The value and benefits of text mining*. JISC <https://www.jisc.ac.uk/reports/value-and-benefits-of-text-mining>
- Neylon, C. (2012). More than just access: Delivering on a network-enabled literature. *PLoS Biology*, 10(10), e1001417. <http://doi.org/10.1371/journal.pbio.1001417>
- Pinfield, S., Cox, A.M. and Rutter, S. (2017). *Mapping the future of the academic library*. SCONUL. Online. <https://www.sconul.ac.uk/news/mapping-the-future-of-academic-libraries> [Accessed May 1, 2018].
- Priem, J. (2013). Scholarship: Beyond the paper. *Nature*, 495(7442), 437-440. <http://doi.org/10.1038/495437a>
- Rao, A. (2017). A strategist's guide to artificial intelligence. *Strategy + business magazine* 87 (1-10).
- The Royal Society. *Machine learning: The power and promise of computers than learn by example*. <https://royalsociety.org/~media/policy/projects/machine-learning/publications/machine-learning-report.pdf>
- Smith, A. (2016). Big Data Technology, Evolving Knowledge Skills and Emerging Roles. *Legal Information Management*, 16(4), 219-224.
- Tredinnick, L. (2017). Artificial intelligence and professional roles. *Business Information Review*, 34(1), 37-41.
- Underwood, J., & Luckin, R. (2011). What is AIED and why does Education need it. A report for the UK TLRP TEL-AIED Theme.
- Vincze, J. (2017). Virtual reference librarians (Chatbots). *Library Hi Tech News*, 34(4), 5-8.
- Woods, D.A. and Evans, D.J. (2018). Librarians' Perceptions of Artificial Intelligence and Its Potential Impact on the Profession. *Computers in libraries* 38 (1) 26-30.