Monograph:
Cox, A. orcid.org/0000-0002-2587-245X (2021) The impact of AI, machine learning, automation and robotics on the information professions: a report for CILIP. Report. CILIP

The impact of AI, machine learning, automation and robotics on the information professions

A report for CILIP

CILIP The library and information association

Research report: The impact of AI, machine learning, automation and robotics on the information professions

This report is published by CILIP: the library and information association with the support of Health Education England.

It was written by Andrew M. Cox (Information School, University of Sheffield).

With grateful acknowledgement to Sue Lacey Bryant, National Lead NHS Library and Knowledge Services for her support and guidance in the development of this report.

About CILIP

CILIP is the UK library and information association. We are the only chartered body in the world dedicated to uniting, supporting and advocating for information professionals and librarians – the people who help the world make better decisions. Our membership is open to everyone working in libraries, information or knowledge management, data science and analytics or a related professional role. We work with employers, learning providers and suppliers across the library and information sector in the UK and internationally to develop talent, promote innovation, encourage workforce diversity and ultimately to secure the long-term future of our profession.

Reference group

The author would like to extend this thanks to the following colleagues who generously took time to share their insights with the project:

Stephen Ayre (William Harvey Library, George Eliot Hospital), Ken Chad (Ken Chad Consulting), Alex Fenlon (Library Services, The University of Birmingham), Julie Glanville (Independent consultant in Information Retrieval), Felix Greaves (Director, Science, Evidence and Analytics at National Institute for Health and Care Excellence (NICE)), Jake Hearn, Timothy Jacobs (The Christie NHS Foundation Trust), Indra Joshi (Director AI, NHS), Niel Kempson, Tony Lewis (TmL Consultancy), Stephen Phillips (vice chair CILIP K&IM committee), Alexandra Pooley (Chair SLSIG), Mia Ridge (Digital Curator, Western Heritage Collections, Digital Research team, British Library), Hélène Russell (The Knowledge Business), Tony Russell-Rose (Goldsmiths, University of London), Edward Saperia (Newspeak House), Philip Scott (Reader in Health Informatics, School of Computing, University of Portsmouth), Matthew Soare (Makerspace Development Manager, Hull Central Library), Alex Smith (Global RAVN Product Lead), Marion Spring (Associate Director, Information Services, National Institute for Health and Care Excellence (NICE)), Martin White (Managing Director, Intranet Focus Ltd).

Published in May 2021.

Contents

03 Executive summary

05 The context and the brief

The context
The brief

07 CHAPTER 1 **Automation in the 4th Industrial Revolution: grasping the complexity**

12 CHAPTER 2 **The promise, risks and threat of AI and robots**

2.1 The promise of AI and robots
2.2 The ethics of AI
2.3 The threats to the profession from AI and robots

17 CHAPTER 3 **Applications of AI and robots in information work**

3.1 Intelligent web and mobile search: the need for algorithmic literacy
3.2 AI interfaces to existing knowledge discovery systems
3.3 AI in knowledge discovery: the need for socio-technical infrastructures
CASE STUDY Living systematic reviews
CASE STUDY Library research collections and the creation of a community of data scientists

3.4 AI interacting with users: new ways to connect
3.4.1 Conversational agents aka chatbots
3.4.2 Voice assistants

3.5 AI managing people: questions about ethical use
3.5.1 User management
3.5.2 Human resource management
3.5.3 Marketing

3.6 Robotic process automation (RPA)
CASE STUDY The ad hoc application of AI to a specific information challenge

3.7 Smart spaces and robotics
3.7.1 Smart spaces
3.7.2 Robots and libraries

3.8 4th Industrial Revolution automation: the need for AI, robotics and data literacy
CASE STUDY Create Crates

27 CHAPTER 4 **The challenges and opportunities**

4.1 The hype challenge: the need for understanding and a vision

4.2 The ownership challenge: the need for collaboration

4.3 The procurement challenge: the need for expertise in licensing, copyright and IPR

4.4 The data challenges: the need for data stewardship

4.5 The technical infrastructure and workflow challenges: the need for information architecting

4.6 The post-implementation challenges: the need for support, training and promotion

4.7 The ethical use challenge: the need for responsible use rooted in professional values and ethics

33 CHAPTER 5 **Competencies needed to take the opportunities presented by AI and robots**

5.1 The relevance of existing information professional competencies

5.2 Computational thinking, fusion skills and soft skills

5.3 Data science and data stewardship

39 CHAPTER 6 **Strengths and vulnerabilities of the profession**

42 **Recommendations**

For CILIP

For information services and libraries

For individual information professionals

For educational institutions at all levels

For other training providers

For research

44 References

48 Appendixes

52 Resources and case study sources

Executive summary

The purpose of this research is to help CILIP and our professional community to understand how AI, machine learning, process automation and robotics are either already impacting the daily work of information professionals or likely to do so in the near future.

Specifically, it seeks to answer the following questions:

1. **How do we ensure that today's workforce has the skills and understanding they need in order to enable them to support their users in participating safely and successfully in a modern world that is increasingly powered by artificial intelligence (AI), machine learning, process automation and robotics?**
2. **What are the ethical implications of our approach to these technologies – how can we deploy the existing ethical framework for librarians and ensure that it aligns to emerging work on Data Ethics and responsible technology?**
3. **What should the skillset of the future workforce look like and what is the curriculum by which we will ensure that the next generation of information professionals have the skills to keep pace with future developments in technology?**

The primary function of the outputs of the research will be to drive the refresh of CILIP's *Professional Knowledge and Skills Base*, and from this the "ecosystem" of teaching, learning and professional development for current and future information, knowledge management and library professionals.

The data for the project was an extensive literature review and conversations with twenty-one experts from across the UK profession conducted by the author in November and December 2020.

In the interests of direct expression the phrase "AI and robots" is used in the report to define the scope of technologies under consideration, although there was little consensus about terminology in the literature or among interviewees.

For many reasons this is a complex topic: partly because of the lack of agreed terminology, but also because of the emotive responses it evokes, the sectoral patchiness of its impact, the complex link between AI and the wider digital transformation, and its interrelations with other social changes. The complexity of the issues is summarised in **chapter 1**.

Chapter 2 provides an overview of how AI and robots are relevant to the work of the profession. This arises particularly from how it might change how text and other forms of content are described, searched and used. This could constitute a shift from the paradigm of searching for items to read manually to the model of extracting knowledge from collections of text, images and other data through algorithms. But there are changes happening across the whole information value chain: in its production, organisation and consumption. The technologies open up exciting opportunities for a more information rich world. However, they also pose a large number of ethical concerns. They will also have a potential impact on jobs, including on the work of information professionals.

The most obviously relevant change AI has effected is in web and mobile search, so **chapter 3**, which identifies the applications of AI and robots in information, knowledge management and library work, starts with this. Changes to web and mobile search imply a need for what has been called algorithmic literacy, as an aspect of information literacy. AI will also increasingly appear in interfaces to knowledge discovery systems licensed or run by information services. However, it is in its impact on knowledge discovery that AI will have the greatest impact. What types of socio-technical infrastructure are needed to support this and how information professionals might be involved varies a lot between contexts. Conversational agents or chatbots offer a new means to communicate with users. AI might also be cautiously used in managing people, but the ethical issues here are significant. Automation of routine administrative tasks including within AI pipelines is facilitated through robotic process automation. Sensor data could be used to make buildings such as libraries smarter to improve design and for wayfinding. Four potential uses of robots in libraries are in handling books, acting as an information desk, for learning, and for generic tasks such as cleaning or through robotic furniture. Finally, there is a potential library role in promoting public understanding of AI and robots, through AI and data literacy. Case studies provide more in-depth narratives around some of the key applications and how information professionals can be involved.

Chapter 4 suggests how opportunities for information professionals are created through the need to overcome organisational challenges to reaping the benefits of AI and robots, especially in knowledge discovery. These include a vision; collaboration; expertise in licensing, copyright and IPR; data stewardship; information architecture; support, training and promotion; and responsible use rooted in professional values and ethics.

Many of the competencies needed to take these opportunities already exist at some level in the profession and are reflected in the CILIP Professional Knowledge and Skills Base (PKSB). **Chapter 5** presents these. A longer discussion is offered about the repositioning of the profession in relation to computational thinking, data analysis and data science, the fusion skills to manage the relations between humans and AI, and the soft skills that AI are unlikely to develop in the near future. A concept of data stewardship is also defined.

The information, knowledge management and library professional is in a strong position to help organisations move into the 4th Industrial Revolution deploying AI and robots, which is reiterated at the start of **chapter 6**. Yet some vulnerabilities remain, particularly in relation to adjacent professions. These vulnerabilities can be partly addressed through the **thirteen recommendations** the report makes to six different stakeholder groups. Of particular importance are:

- For CILIP to articulate the relevance of the profession's skills, values and ethical principles, to identify pathfinder organisations and people who encapsulate these possibilities, and to promote knowledge sharing within the profession.
- For individual information professionals and information organisations to explore the use of AI tools and share what they learn with others across the profession.
- For the profession to use the technologies in their own work and to support information users in engaging productively and safely with them for social good.

> **For many reasons this is a complex topic: partly because of the lack of agreed terminology, but also because of the emotive responses it evokes.**

The context & the brief

by Nick Poole, CILIP CEO (November 2020)

The context:

We are in the middle of a 4th Industrial Revolution, every bit as transformative as the great industrial and economic shifts that preceded it.

Building on the great advances in information technology and connectivity in the postwar era, this new revolution is described by HM Government in their White Paper Regulation for the Fourth Industrial Revolution:

"The Fourth Industrial Revolution is of a scale, speed and complexity that is unprecedented. It is characterised by a fusion of technologies – such as artificial intelligence, gene editing and advanced robotics – that is blurring the lines between the physical, digital and biological worlds. It will disrupt nearly every industry in every country, creating new opportunities and challenges for people, places and businesses to which we must respond."

The White Paper envisages four 'grand challenges' for the UK economy and industry if we are to position ourselves to take advantage of these changes:

1. **To put the UK at the forefront of the artificial intelligence and data revolution**
2. **Maximise the advantages for UK industry of the shift to 'clean' growth**
3. **Become a world leader in shaping the future of mobility**
4. **Harness the power of innovation to help meet the needs of an ageing society**

These challenges have a profound implication for the future development of the information, knowledge management and library workforce. The pace of adoption, the ubiquity of the associated technologies, the fundamental changes in user behaviour and the sheer scale of machine-readable information dwarf the challenges of structured search on the Internet and the digitisation of print books that characterised the late 20th century for our profession.

As is so often the case with these large-scale industrial changes, the most visible part is technological. We can 'see' the mainstreaming of mobile Internet-connected devices, the use of smart technology in the home, the emergence of AI-powered assistants in the interfaces we use in our daily lives, the near-driverless cars, etc.

Harder to discern but arguably much more significant are the huge changes in human and physical infrastructure that enable these innovations – the data centres built in the desert powered by hydro-electric dams, the expansion of network cables and the emergence of a workforce with the skills needed to make it all work.

The figures go some way to illuminating the scale of social, economic and industrial re-engineering that has been going on for much of the past decade (source: KPMG research data for CILIP/KPMG Report Information as an Asset):

- Chief Data Officers in large enterprises had an average budget in 2017 of £6m, a 23% increase year-on-year since 2014;
- 10% of major corporations in 2017 predicted they would have a 'Chief Robotics Officer' in their senior executive team by 2021;
- The total volume of data being created, managed and shared globally is predicted to increase tenfold between 2016 and 2030 (from 16Zb in 2016 to 162Zb in 2030);
- In 2020, 1.7Mb of data are being created per person, per hour globally;
- The global market for AI and machine learning technologies is predicted to grow from £16bn in 2018 to £143bn in 2025;

- The global market for industrial and non-industrial robotics is predicted to grow from £22bn in 2018 to £178bn in 2025;
- The number of 'Internet of Things' smart connected devices is predicted to grow from 27bn in 2017 to more than 125bn in 2030.

Every day, the technologies associated with the 4th Industrial Revolution – AI, machine learning, automation and robotics – are finding new applications. Examples include:

- Fraud/risk detection and management in financial services
- Regulatory compliance at scale (for example with Data Protection legislation)
- Decision support
- Automation of inventory across supply-chain, logistics and stock management
- Use of chatbots to automate customer service enquiries
- Use of self-help apps, for example for self-management of long-term conditions

The brief:

The role of the information, knowledge management and library professions.

The tremendous scale and pace of the 4th Industrial Revolution and its associated technologies will inevitably disrupt the library, information and knowledge management workforce. In 2017, McKinsey estimated that 50% of current work activities were technically automatable, and that 60% of jobs have at least 30% of their responsibilities that are capable of being automated.

The opportunities are potentially immense – people will lead increasingly information-rich lives in which they can use data and analytics to improve their health, be more productive and release more time for family, friends and leisure activities. In so doing, they will need expert, ethical and accountable information professionals to help them select and use the right technologies, to maximise their information-handling skills and to keep them and their families safe from harm.

With these opportunities come considerable risks– for example, a specialist UCL research unit has already found examples of Internet-connected smart home devices being used for coercive control in the home by abusive partners. The literature and news are full of many examples of unaccountable algorithms and AI-driven decisions which exacerbate human failings such as sexism and prejudice.

The role of information professionals has always been to enable our users and communities to profit from new advances in technology and to make better use of information for their own advancement. This was as true of the cuneiform collections of antiquity as it was of the great book and print collections of the past century as it will be of the vast collections of data and personal information being generated and shared online.

The opportunity and the challenge for CILIP as a professional association is to ensure that we are able to use our 'pipeline of talent' to prepare the information professional workforce to play this role once again in the 4th Industrial Revolution.

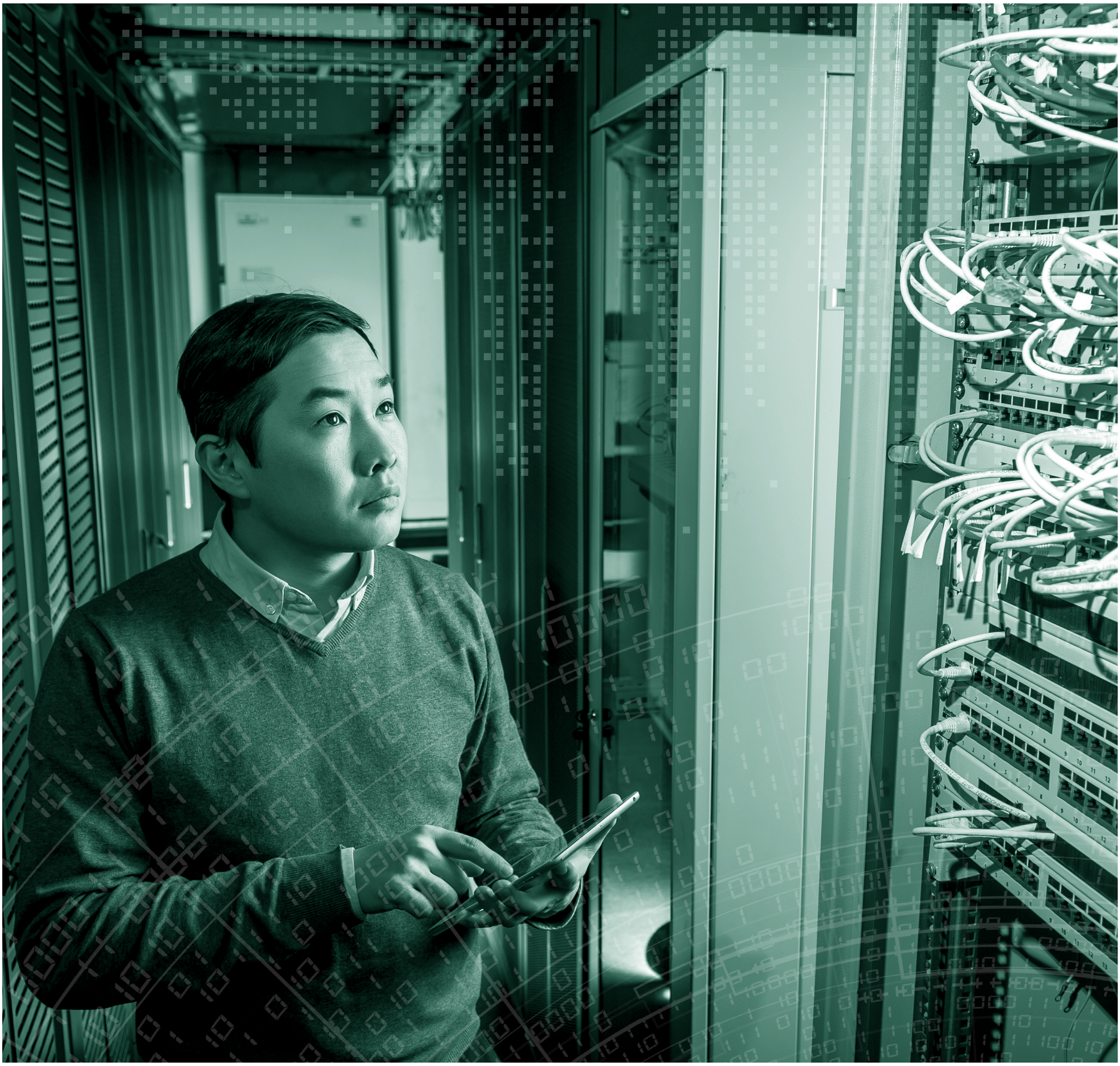
Specifically, there are three key questions which we will need to be able to answer in order to ensure that our profession is ready and able to fulfil this role:

- How do we ensure that today's workforce has the skills and understanding they need in order to enable them to support their users in participating safely and successfully in a modern world that is increasingly powered by AI, machine learning, process automation and robotics?
- What are the ethical implications of our approach to these technologies – how can we deploy the existing ethical framework for librarians and ensure that it aligns to emerging work on data ethics and responsible technology?
- What should the skillset of the future workforce look like and what is the curriculum by which we will ensure that the next generation of information professionals has the skills to keep pace with future developments in technology?

> **"The scale and pace of the 4th Industrial Revolution will inevitably disrupt the library, information & knowledge management workforce."**

1

Automation in the 4th industrial revolution: grasping the complexity



1 Automation in the 4th industrial revolution: grasping the complexity

There are many reasons why the pattern and implications of the current wave of automation are hard to fully grasp.

AI and robots are our past, present and future.

When will automation impact us? All the technologies behind AI and robots have a long history. There seems to be a tendency to treat AI applications that have already been implemented as not counting as AI. For example, many of the familiar features of Google and Amazon are powered by AI. Digitisation techniques such as OCR are a form of machine learning. RFID is an instance of the Internet of Things. The introduction of book returns and sorting machines in physical libraries is effectively an application of robotics. This report is not about these technologies, but it is relevant to remember how the process of implementation happened and ask how the implementation of the latest round of automation might be the same or different.

While the roots of current changes may lie in the past, it is also important to remember that the diffusion of innovation takes time. Proof that a concept is potentially valuable through special projects does not always lead quickly to organisation or sector wide change or necessarily to change at all. For example, chatbots are widely used in some industries and have been promoted for a while as relevant to information services, such as libraries. There does not seem strong evidence that the uptake has been what might have been predicted. This could be because of the nature of reference enquiries being too complex or the continuing value of human contact. Self-issue and returns systems spread more quickly, but it did take years for it to become business as usual, starting with the first automation initiatives in the 1930s right through to automation becoming “business as usual” in the 1990s.

While it is true that AI, robots, even machine learning have a long history, and some promises of technology never materialise, at the same time it is also true that things are changing rather fast and things that seemed impossible a few years ago are being achieved. Some of these changes could constitute a disruptive change. They could vastly improve information access; they might change information professional jobs and indeed in some sectors such as health, they already are.

AI and robots are ideas rather than specific technologies.

If they were specific technologies their precise impact might arguably be relatively easy to delimit. But in fact they are both ideas: ideas about machines emulating or even replacing humans. AI imagines machines that can reason, make decisions, learn and interact like humans. Robots mimic or substitute for human bodies. The idea of a machine that can think or has a body has been around a long time and there is a strong public imaginary of what they might mean, which we see expressed through the media and popular science fiction movies. The notion of automation also has a long history. The figure of the handloom weaver swept away by the first industrial revolution echoes in our minds. The long cultural history attached to ideas around automation, along with the spectre of job losses, complicates our understanding of what is happening right now.

AI and robots produce emotive responses.

Are AI and robots a "good thing"? The immediate arrival of the technologies is surrounded by excitement and some hype, secrecy and fear. AI promises to make information more accessible, through indexing material more quickly and in new ways, such as by permitting searching by voice command. Excitement is well justified, particularly for information professionals whose core value is that everyone has access to the information they need.

Yet in some very competitive sectors AI is being used as marketing, to present organisations as innovative and on the cutting edge. For the same reason precisely how it is being used is kept secret. Sometimes this may mask that there is little that is being offered that is new. At the same time much media coverage focuses on the fear that AI and robots will steal people's jobs and make bad, biased decisions. High profile failures such as of the algorithm used to calibrate A level results in the summer of 2020 cast suspicion on the technology. The emotion and secrecy is not very helpful for getting a clear picture of how the technologies will impact our profession's work. For information professionals working in such highly competitive sectors this could be a particular challenge. In other, less competitive, more open sectors the new technologies can be appraised in a cooler light. Yet in all cases AI does carry with it a host of ethical issues.

AI and robot adoption is patchy geographically and by sector.

Where are AI and robots being applied? Another complication making it hard to understand patterns of change is that the impact of technologies arrives in different sectors of work and different parts of the world at a differential rate. What is futuristic in one industry or in one country has already happened in another. Many of the key aspects of AI are already quite familiar in technologies such as in web and mobile search (auto-suggest, relevance ranking, recommendation, personalisation) or word processing (auto-complete, spell checking, translation, voice recognition). However, they have yet to arrive in most search systems used in information services. This carries a danger of making the specialist tools we support looking outdated.

On the whole, we can guess that most AI will be developed for application in high value tasks and then at a later date applied in information contexts, such as libraries. This means that what is being adapted to our context may often make assumptions that do not fit our usual way of doing things or the language we use to describe things. This pattern of development also potentially limits our ability to shape the technology.

Reflecting the patchiness in change, because the information sector is so broad and the roles we carry through quite diverse, when and how AI and robots impact on the information profession will also be patchy. Specifically, the differing levels of digital maturity across sectors will shape their impact. How AI might impact a research library will be quite different from how it impacts a legal information professional, a knowledge and information service in government or the health service, or an information professional working in a bank. Large organisations are more likely to have the resources and data to apply the latest automation technologies, so even within sectors the impact will be patchy.

How information professionals are affected is connected to how AI is impacting their wider sector. The "smart library" could sit within a "smart campus" and beyond that a "smart city", so information related applications are part of a wider nexus of changes, making the nature of the future more unclear. The organisations within which information services operate may be transformed with many effects on information professionals. For example, one thinks of the wide range of applications of AI identified by NHSX (2019) both in clinical decision making (e.g. in radiology and pathology) but also in operational aspects (e.g. managing waiting lists) (see also Stanfill & Marc, 2019). Faggella (2020) sees six main ways that law firms are applying AI to their operations: in due diligence, predicting litigation outcomes, legal analytics, analysis of IP portfolios, automation of document completion, and billing for partners' time (see also Kroski, 2020; Neary & Chen, 2017; Smith, 2016). Some of these directly impact on traditional areas of information work; others may have ripple effects on them. In the university

> **Most AI will likely be developed for high value applications, then later applied in information contexts – potentially limiting our ability to shape the technology.**

context, Cox (2021) identifies a wide range of applications in teaching (Intelligent Tutoring Systems, conversational agents in MOOCs, plagiarism detection), research (text and data mining, robot scientists, automated peer review) and university administration and estates management (the smart campus). This means that information professionals embedded in their sector may see very different patterns of change happening around them. Information professionals work in every sector so a comprehensive perspective is impossible.

AI is an umbrella term, not a single technology.

What is AI? Terms such as machine learning, natural language processing, voice recognition and computer vision refer to some of the computational techniques associated with the idea of AI. The next section explores in more depth the range of technologies involved, but it is important to say that each has its own history and impacts. It might have been easier to delimit the topic of this report if we had focussed just on machine learning, say. But the wider scope set is productive in reflecting that there are a range of technologies and interrelated developments involved. At the same time AI is premised not just on clever algorithms, but also on modern computing power and on data. Many of the algorithms were written some time ago; it is only now that the computing power exists to execute them fast enough to create usable services from them. And, it is only now that there is enough data to allow them to be useful.

There is not yet agreed terminology. Each of the technologies behind AI has its own history. None are wholly new. But sometimes this is partly masked by shifts of terminology. A few years ago there was a lot of talk in parts of the profession about text and data mining (TDM), linked to new exception to copyright creating a right to mine texts. Now we would probably talk much more about AI or machine learning. This makes tracing the growing impact of AI a little more difficult.

There is not yet agreed terminology. Each of the technologies behind AI has its own history, but none are wholly new.

As is typical of an arena of change, terminology has not settled down. Some interviewees for this report, for example, felt very comfortable with the term artificial intelligence; others thought it unhelpful, because of its connotations or lack of agreed definition. Yet other terms, such as machine learning, although an important reference point, are too narrow. For the purposes of the report we will prefer the term AI because it usefully captures the range of technologies supporting 4th Industrial Revolution automation. The term AI is often used in the literature to encompass robots, but sometimes it will be useful to differentiate AI and robots.

There is also a blurred boundary between the bundle of technologies addressed in the report and others, such as virtual reality (VR) and augmented reality (AR) that are also making their own way through the famous Gartner hype cycle. Perhaps their application will be interlinked with AI; several authors suggest so. But this does make it very hard to delimit the report. We have decided to exclude VR, AR, and blockchain from direct consideration here, but this may mean that some trends are missed.

Even within the technologies often associated with AI there is significant difference. For example, robotic process automation (RPA) automates routine tasks. It increases speed and reliability of carrying through routine tasks. The argument for using RPA relate to cost saving. In contrast many of the claims behind artificial intelligence promise to do familiar things in new types of way. Their allure lies in the promise a true digital transformation and disruptive change.

AI and robots will create, change and destroy jobs.

What will the impact of AI and robots be on jobs and skills? AI will replace human work, particularly repetitive forms of knowledge and service work. Frey and Osborne (2017) predict the disappearance of many jobs, including "clerical library assistants" and "library technicians" (though these roles are not clearly defined). Robots will also replace human labour, again particularly in highly repetitive, predictable tasks. Similarly, McKinsey (2018) suggest that jobs based on both basic cognitive skills and physical and manual skills will shrink. This change could be in the form of automating the most repetitive tasks, allowing the employee to focus on the more interesting aspects of jobs, such as those requiring imagination, problem solving and collaboration.

It could lead to role expansion. Work could be rehumanized (Wendehorst, 2020). But it could lead to de-skilling too, a more polarized workforce or perhaps to employee disempowerment (Coombs et al., 2020; Wendehorst, 2020).

AI and robots will also create many jobs: people developing tools, analysis, and for domain experts. Indeed, for the foreseeable future most AI will keep a "human in the loop" to ensure that human judgement can shape critical decisions. Many experts on AI differentiate weak and strong AI, and see weak AI as a near reality in a way strong AI is not. Weak AI does a specific task very well and probably more accurately and faster than any human. Strong AI that applies a general understanding to different contexts is much harder to achieve, perhaps even impossible. It is certainly many years off. The same may be true of robotics, where robots that perform very narrow specific tasks, such as sweeping the floor, are far more likely than ones that can perform a range of tasks as in our idea of a human-looking domestic robot. So McKinsey (2018) anticipate the greatest growth in technological skills. They also predict a growth in the need for higher cognitive skills, though they see a decline in the value of "advanced literacy and writing" which may be significant for some information professional work. But they also see a shift in increased emphasis on social and emotional skills, areas where automation is much harder. Thus, the balance in our skillsets may be shifting.

AI and robots are part of the digital transformation.

Where do AI and robots fit into wider changes? A further complexity is that we need to reflect on how the current wave of technologies fit into the wider digital transformation. Is it a step change in the digital shift or just one more step on the move to digital? If the latter we might need to be thinking how to adapt our existing models rather than creating wholly new ones.

AI and robots are part of wider social changes.

We also need to factor in other changes to our society that will interact with the new technologies. This includes both dramatic events such as Brexit and the COVID-19 pandemic, and more long-term social trends such as ageing population structures or a growing concern with environmental sustainability. These wider changes do interact with the current wave of technology. For example, one of our interviewees thought that the urgency created by COVID-19 has pushed ahead innovation in the automation of systematic reviewing faster than might have been expected. The Black Lives Matter movement seems to have helped change the moral climate within which bias in AI systems is being challenged. And although often presented as entirely immaterial, as virtual, AI does have significant environmental impacts through its demands on the power system. It also creates potentially exploitative relations, e.g. for clickworkers. Untangling such complex nexuses of change is beyond the scope of this report but is relevant to how the reader needs to use its findings.

The future with AI and robots is not predetermined.

Do we have power to influence how AI and robots are used? A final point to emphasise is about the nature of technological change. Sometimes technologies are presented as external forces that cause inexorable changes to our lives as if from outside. This form of technological determinism does reflect how we sometimes experience change, particularly when it is badly managed. But in truth technologies are social creations. Ultimately it is societies that make choices about technologies they develop and use. By extension the information, knowledge management and library workforce has opportunities and an important role in choosing about the collective future through how it responds to the current wave of technologies.

> **AI will replace repetitive forms of knowledge and service work, and robots will replace highly repetitive, predictable labour.**

2

The promise, risks and threat of AI and robots

2 The promise, risks and threat of AI and robots

2.1 The promise of AI and robots

AI applications are already with us, most obviously in web and mobile search, in the form of features such as ranking of results, auto-suggestion and personalised results. For everyday users, these applications are experienced as helpful, informative and fun. AI is used to offer recommendations in ecommerce sites. In daily information tasks such as writing a document we benefit from auto-correction. Automated Writing Evaluation tools (such as Grammarly) are increasingly being advertised to students and others. A little more novel are auto-summarisation of text, auto captioning of videos (through speech recognition) and translation of documents. While not working perfectly, these new tools make information tasks easier and quicker.

From these examples we can understand Gartner's (Eliot et al., 2020: 5) description of AI as a "general-purpose technology" and their claim that "AI will be in everything."

AI is often used as an umbrella term for a number of technologies. Thus Gartner (Lowendahl & Calhoun Williams, 2018) refer to "six core interconnected AI technologies": business analytics and data science; natural language processing, speech recognition and text to speech; machine learning, deep learning and neural networks; machine reasoning, decision making and algorithms; computer vision; and robots and sensors.

Similarly, McKinsey (2019) identify a number of inter-related strands of "AI capability" in: robotic process automation; computer vision; machine learning; natural language text understanding; virtual agents or conversational interfaces; physical robotics; natural language speech understanding; natural language generation; and autonomous vehicles.

The changes in how text can be processed are most likely to impact information professional work most strongly.

In both these definitions AI is seen as encompassing a wide range of computational techniques and includes different forms of robotics.

AI has been defined as: "computer programs that perform tasks requiring intelligence when done by humans." (cited Holmes et al., 2019)

A contrasting definition is:

"An AI system is a machine-based system that is capable of influencing the Environment by making recommendations, predictions or decisions for a given set of Objectives. It does so by utilizing machine and/or human-based inputs/data to: i) perceive real and/or virtual environments; ii) abstract such perceptions into models manually or automatically; and iii) use Model Interpretations to formulate options for outcomes". (OECD, 2019)

The computational techniques that make up AI differ in importance in different settings. For example, in the high tech sector, according to McKinsey (2019), machine learning is the most important technology. In financial services robotic process automation is particularly important; in healthcare services, computer vision and natural language processing. Given that information professionals work in every sector it is not possible to say which of these might be more important for us.

However, it is most likely the changes in how text can be processed that will impact information professional work most strongly, because historically much of our work revolves around text in various forms. Much AI is what we might have called text mining five years ago (Anderson & Craiglow, 2017; Glanville, 2019; Glanville & Wood, 2018; Stuart, 2020). We are used to computers dealing with structured data like spreadsheets and databases. One of the promises of AI is to deal better with "unstructured text" by analysing words to identify concepts and cluster related ideas, whether those texts be published books and journal articles, grey literature, internal know-how or even text harvested from conversations online. Difficult to digitise parts of texts such as tables may be easier to process (Cordell, 2020). AI is improving the digitisation of manuscripts through recognition of hand-written manuscripts. Sentiment analysis even seeks to identify different types of feeling in texts.

The ability of AI to analyse the content of texts may shift search away from being primarily through structured bibliographic databases to search of full text items or whole collections using multiple potential algorithms. The paradigm could be seen as shifting from searching for an item (like a book or online report) to read manually to mining for knowledge in whole collections of text data: from close reading to distant reading. If this is the case, having some grasp of AI techniques will be important to many information professional jobs. Potentially, work information professionals do through creating descriptive metadata might become less necessary as computers perform this task, though this is debatable. The change will not reduce the need to understand the ecologies of the production of information in order to conduct an effective search or the need to critically evaluate results based on this understanding. Yet as Tony Russell-Rose suggested in the interview for this report, AI will impact on the production of texts throughout the information lifecycle or value chain:

- In producing texts. The first scholarly text composed entirely by a computer synthesising existing literature was published by Springer in 2019. Publishers are widely using AI to manage the peer review processes (Thelwall, 2018).
- In acquiring texts into collections, e.g. in crawling for data or processing text
- In categorising texts and knowledge discovery. New tools will exist to describe the content of a single text or a large collection.
- In curating and reuse of text.
- In consuming text. Users will be supported to access text through AI tools such as summarisation which provides them quick access to content or recommendation and personalisation whereby appropriate content is pushed towards them. They may also be enabled to search across whole collections as datasets rather than pages, as Ridge (2019) puts it. Publishers are increasingly using AI for content enrichment, which automates bringing together content from different sources for the reader (67 bricks, 2017).

As well as processing unstructured texts, other forms of information are handled better than before through AI technologies, thus computers are able to:

1. Process inputs in the form of human voice (e.g. voice search) and also to output in voice form.
2. Identify the subject of images and categorise them better, as in a search engine being able to identify what an image is of or enable searching using an image rather than a text term, or as in facial recognition, where it is able to match a human face to a photograph. Computer vision is the area of computing which deals with the extraction of information from images and video.

This means that computers can accept a wider range of inputs as data, including images, sound recordings and videos. Natural Language Processing (NLP) is the term for turning text or audio into structured information. They can also interact with people to receive and give data through voice or images. Chatbots can draw on a knowledge base to interact with a human in a more or less human interactional style.

Machine learning means that computers can be given training data from which they can identify useful patterns, without those patterns necessarily being determined by a human in advance. Supervised machine learning is where training data that has been manually labelled is used to help the computer to learn to recognise patterns. In unsupervised machine learning, the algorithm searches for patterns without being given training data.

There is nothing new to automating business processes, but what robotic process automation (RPA) offers is relatively user-friendly tools to stitch together processes across different platforms. This promises to make it easier to automate daily processes.

These abilities widen the range of data that computers can accept, processes they can carry through and outputs they

Having some grasp of AI techniques will be important to many information professional jobs.

Biases in training data (and in the computing industry itself) systematically reinforce existing unfair social structures.

can generate. It allows them to do things in some sense akin to human thinking, learning, decision making and interacting.

The ability to train AI rests on the availability of data. AI is premised on "big data". As the points above illustrate, the range of inputs and outputs that computers can handle has expanded to unstructured text, voice and images. In this sense the definition of "data" has expanded. What librarians might call collections and more latterly might have been referred to as "content" can now be seen as data (Padilla et al., 2018). Books and journal articles are data, as are grey literature, know-how, photo collections and archives of email or online forums. This is not just an important shift in language. This is a critical change because many professional distinctions rest on differentiating types of content, such as between published books and articles in libraries, unique manuscripts in archives and historical objects in museums. Increasingly the boundaries blur and issues converge.

In addition, the digital shift has led to a growing range of other forms of computer-accessible data. Online activities leave data traces. Sensors embedded in objects and buildings gather more data about the world, including ourselves. We carry devices such as smartphones that track our own activity, location and movements. Indeed, even human feeling is claimed to be data for affective computing. This is associated with a societal re-evaluation of data, with claims that data is the "new oil". Actually, this claim was first made as early as 2005 by the mathematician involved in developing Tesco's supermarket reward scheme, the Clubcard. His point was that oil is valuable, but needs to be refined. Similarly, issues of data quality, of data description, data management and governance apply in the AI context.

This is in many ways very exciting from an information, knowledge management and library perspective, because an AI-driven and data-intensive society is one based on information and knowledge. The amount of data available to humanity is growing and in many contexts AI is making it is easier to find information. The potential to search across different formats of information is increasing. Computers are better at supporting our understanding through interacting with us orally and presenting new visualisations of data.

Equally the need for a critical awareness about how knowledge is generated grows, and the information, knowledge management and library profession has an important part to play in this.

Nevertheless, it is important to recognise that in many contexts these developments remain potential. The admittedly fragmentary evidence suggests a great interest in the technologies in all sorts of organisations in the UK (AIIM Industry Watch, 2018; British Computing Society, 2020; Carter, 2019). But as section 3 made clear, different sectors are at different stages of acceptance of the technologies.

2.2 The ethics of AI

In addition to questions about the maturity of the technology, there is intense concern at the societal level around the ethics of AI. Jobin et al (2019) report on no less than 84 statements attempting to define what constitutes ethical AI, including by the House of Lords, the European Union and the OECD; the Royal Society and IEEE; as well as Google, IBM and Microsoft. The timeline in Fjeld et al. (2020) suggests that these concerns emerged in 2018 and 2019. So the impact of AI is surrounded by a storm of controversy in society as a whole (Centre for Data Ethics and Innovation, 2020; Equality Task Force, 2020; Floridi & Cowls, 2019) as well as within the information profession itself (Association of Research Libraries, 2019; Cordell, 2020; IFLA, 2020; Padilla, 2019). Despite this evidence of concern, there is more to say because most of these statements originate from the global North and questions remain about the implementation of these codes in practice (AINow, 2019). The following paragraphs summarise some of the main concerns.

If the data used to train an algorithm is based on a biased sample then it will learn a misleading pattern. If the data we train computers on only reflects our past faulty decisions, then it will just repeat our past biases. Computers trained on white faces do not necessarily identify black faces. This raises issues of *fairness and non-discrimination*. Much of the literature points to the way that these biases in training data (and in the computing industry itself) systematically reinforce existing unfair social structures.

Privacy is another concern. Because these technologies are embedded in other tools and sites as general purpose technologies, the user may not be aware that they are gathering data about them, trying to influence or nudge their behaviour or making decisions based on it. This raises issues of consent, the right to control over how one's data is used and the right to be forgotten.

In so far as we do have a sense that data is being collected, AI can be creating surveillance. Quite apart from our right to privacy, this awareness of surveillance can in itself change our behaviour, e.g. make us reluctant to search for particular forms of information. Surveillance impacts the right to free expression.

Gathering data about us creates a *security* risk, if it is compromised or stolen and appropriated.

The operation of AI in personalising content and supporting moderation has direct impact both on the *freedom to form an opinion* and on the right to *free expression* (Kay, 2018).

Because the technologies "learn" the patterns it may not be easy for anyone, including system designers, to fully understand and explain the outputs or decisions they make. This opacity leads to problems of *explainability and transparency*. How do we explain how the outputs of AI were arrived at? Who is accountable for a mistake if a computer makes it? How can an appeal be lodged? The shifting nature of machine learning may also reduce the possibility of *verifiability and replicability*.

There are also fundamental issues around the *human control over technology, professional responsibility and the promotion of human values* (Fjeld et al., 2020).

These issues are particularly acute where AI is used to make decisions that directly shape people's lives, influence or "nudge" human behaviour, or wherever personal data is involved. Indeed, some critics of AI have gone much further in seeing AI as part of a wider movement towards surveillance and the metrification of society, where only things that can be counted are valued.

Ultimately, it is vital that AI and robots are used for social good. The information profession have an important role in weighing up how these ethical issues apply in contexts where they are operating.

A further dimension of the societal impact of AI and robots, is their potential impact on work, including professional work. Section 3 already discussed the hard to predict impact of the technologies on employment. The next section considers the implications for information and knowledge management and library profession.

2.3 The threats to the profession from AI and robots

In addition to the wider issues and social risks, AI and robots pose a number of threats from a professional perspective. One arises if AI can genuinely perform tasks currently carried out by an information professional. For example, if it produces summaries or indexing that are good enough for their context of use, and much more quickly than is possible to do manually. Or if it can efficiently sort books and put them back on a library shelf. This appears to be highly likely with very routine, predictable cognitive or manual work. It may do away with certain more routine work, especially those involving sifting vast amounts of data.

Currently most AI and robot work is narrow. It requires a "human in the loop". Over time it may also prove to be possible to use AI and robots on more complex, unpredictable tasks, including those seemingly based on expert judgement. At this point more skilled tasks may be lost in human jobs. For information professionals this represents another wave of fear of "disintermediation".

There is also a risk that even though they are less fit for purpose, cheap to produce and easy to use results from AI satisfy busy, unskilled users. There is even the risk that that information services are replaced by AI even if the latter do not work so well at all, if ill-informed decision makers fall for some of the hype. Thus, AI implies the restructuring of work, in unpredicted and perhaps disruptive ways.

However, because there are many challenges around AI in practice, there remain opportunities for information and knowledge professionals to contribute. The opportunities can be more clearly understood if we examine the applications of AI and robots relevant to information work.

The information profession have an important role in weighing up how ethical issues around AI apply in contexts where they are operating.

3 Applications of AI and robots in information work

3 Applications of AI and robots in information work

3.1 Intelligent web and mobile search: the need for algorithmic literacy

The first domain where both users and information professionals have already encountered AI is in search engines, through the many novel features of search that have appeared there first, such as auto-suggestion, auto-correction, recommendation and personalisation, and most recently voice search (Fernandez, 2016). Some of these changes have already happened or are continuing to happen. In many ways they are welcomed by the information, knowledge management and library profession as increasing easy and ubiquitous access to information.

Yet with our focus on Information Literacy and as Google sceptics we immediately understand the limits of some of these applications of AI: encapsulated in terms such as keyword appropriation, filter bubbles, echo chambers, and deep fakes.

A critical evaluation of search results requires an understanding of the hidden biases. The term algorithmic literacy has been coined to refer to the deeper understanding of the operations of search engines needed by users to critically evaluate search results (Bakke, 2020; Sander, 2020).

To have algorithmic literacy would be to have an awareness that:

1. Search engines are driven by commercial motives, not simply by a desire to offer access to information in an unbiased way, so that search results are not simply ranked by popularity.

2. There are layers of network bias (Johnson, 2020) in how search results are produced and presented.

3. Some ranking is determined by organisations paying for a better placing.

4. Search engine optimisation is used to promote results in search results.

5. Search engine results are affected by data voids, keyword appropriation, filter bubbles, echo chambers, deep fakes, automated content moderation, etc.

6. Confirmation bias and unconscious bias may impact our own information-seeking behaviours.

This is an area where most information professionals are well informed and already engaged with AI, though they might not be used to calling it that. It is a positive finding of this report that information professionals are at the forefront of promoting better understanding of AI in mobile and web search and many of the same issues apply across the range of AI applications.

3.2 AI interfaces to existing knowledge discovery systems

Just as the first applications of AI that we have probably experienced are in web and mobile search, so the most obvious application is in information service discovery systems. Such tools have existed for some time in third party subscription databases. They may now be offered for library discovery systems. One example application that is being publicised at the time of writing is Yewno, which is being integrated with ExLibris. This offers a new way to explore library content through visualised concept maps (Gramatica & Pickering, 2017).

AI also has applications in supporting the creation of descriptive metadata within conventional knowledge discovery systems (Flannery, 2020) or to identify items to be weeded because of lack of use (Wagstaff & Liu, 2018).

3.3 AI in knowledge discovery: the need for socio-technical infrastructures

Probably the most important application of AI arises from the way it creates potential for information services to support new ways to analyse content. This could be for digital humanities researchers or social scientists to explore special collections in a research library; for lawyers to mine information from a know-how collection; for researchers to mine published literature for new insights in an era of "big publishing"; for government officials mining information sources on an ad hoc basis to support advice to government ministers.

In order for this to happen, challenges around data/ content, choice and training of AI tools, technical infrastructure (e.g. data storage) and workflows, and of forging collaborations and increasing skills have to be solved to create an appropriate socio-technical infrastructure. Here information services have a strong potential part to play.

There are many models of what a relevant socio-technical infrastructure might look like, such as the following (also summarised in table 3.3).

1. A project-specific assembly of digital content and tools to analyse it. The focus here is on creating a temporary infrastructure to accomplish specific project goals. An example could be a digital humanities project in a research library.

2. A local content collection, such as a know-how collection in a law firm. This could be of one type of data or multiple. Bespoke AI tools are trained to analyse this data. The focus here is the unique value from the content.

3. Digitised or born digital content in the special collections of a research library. The focus here is the unique cultural significance of the content, many tools can be applied. Multiple projects work with the collection over time, often in complementary ways. Outputs from the project could enhance the collection, e.g. by creating new descriptive metadata, aiding future discovery.

TABLE 3.3 Alternatives of socio-technical infrastructures to support AI in knowledge discovery

		Content	Technical infrastructure	Information service role
1	A projects-pecific assembly of digital content and tools.	Temporary, researcher chosen.	Temporary, ad-hoc.	Background support.
2	A local content collection, such as a know-how collection in a law firm.	Locally developed, unpublished content.	Permanent.	Assembling content, training tools, supporting usage.
3	Digitised or born digital content in the special collections of a research library.	Local unpublished content, often digitized.	Temporary for each project.	Providing content, supporting usage.
4	A sector wide aggregation of content available for mining.	Published content.	Collective.	Signposting the service to users.
5	A tool/ content agnostic infrastructure within which bespoke collections of content for different types of analysis are stored.	Any.	For storage and tools.	Building and promoting infrastructure.
6	A publisher or aggregator platform including content and tools – subscribed to by an information service for users.	Publisher owned or aggregated.	Licensed.	Licensing service and supporting use.
7	A semi-permanent assembly of content, infrastructure, tools and people for a very specific, ongoing purpose, such as in a living systematic review.	Information service selected from published sources.	Semi-permanent.	Controlling an integrated solution.
8	A support service that focuses on advice on tools, training and supporting collaboration.	None.	Human rather than technical.	Assembling expertise and building community.

CASE STUDY

Living systematic reviews

This case study illustrates the issues around model 7, in table 3.3. In the health sector (and increasingly in other areas such as education and agriculture) systematic reviews are an important information activity. Because they are critical to guiding clinicians on healthcare, but because of the high rate of publication in this field, the use of AI is increasingly attractive (Beller et al., 2018; Elliott et al., 2014; Grames et al., 2019; Jonnalagadda et al., 2015; Millard et al., 2018). So long as the process and algorithms in use could be clearly defined there would also be increased transparency and reproducibility, aspects which are important in a health context.

Systematic reviews are generally considered to have the following steps:

1. Identifying that a review is needed or an existing one needs updating
2. Identifying a search strategy, including keyword and database selection
3. Screening for items that meet inclusion criteria
4. Collecting the full text of results
5. Evaluating the items found, including assessing risk of bias in trial results
6. Writing the synthesis

The concept of the living systematic review (a self-driving SR) encapsulates the ambition to automate the whole process, so that the review can always be up to date at the level of each guideline.

Many of the individual steps in the process have been approached with a view to automating them, with some, such as collecting full text or supporting human screening, being easiest to achieve. Tools have also been developed to assess risk of bias in trials or to identify sample sizes in reviewed papers. Views differ on time frame within which the various steps in the process can be automated. It is likely that there will be a human in the loop for the foreseeable future, and so they might be better referred to as computer assisted reviews (O'Connor et al., 2019). There remain barriers to acceptance (O'Connor et al., 2019), in funding of development, managing crowd sourcing and licensing issues. There are also critical ethical issues in terms of ensuring transparency and considering the issue of liability if an error is made based on a computer decision.

In the conventional systematic review, the information professional plays a critical role at the stage of understanding the topic, choosing databases to search, defining a precise search to ensure all relevant papers are found, and perhaps in retrieving results. One of the most achievable applications of AI is to help screen hits, by allowing a user to train it to sift a large number of potential hits rapidly. This means that techniques of precise searching become less relevant. This is exactly where many information professionals have tended to fit into the process, so the effect could be to diminish their role significantly. Yet they may be able to contribute to areas such as licensing, creating workflows, training and support.

4. A sector wide aggregation of content available for mining (as planned at one time by JISC). Here the content is published rather than local or unique content. The focus here is primarily on the content with some ready-made tools supplied.

5. A tool/ content agnostic infrastructure within which bespoke collections of content for different types of analysis are stored. The focus here is on providing a technical infrastructure, which users can use in different ways.

6. A publisher or aggregator platform including content and tools – subscribed to by an information service for users. This offers an integrated solution, but at a high cost and loss of researcher independence in choosing content and tools.

7. A semi-permanent assembly of content, infrastructure, tools and people for a very specific, ongoing purpose. This is a complete solution assembled for a specific organizational purpose. A good example would be a living systematic review.

8. A support service that focuses on advice on tools, training to use the tools and supporting collaboration, but does not provide a technical infrastructure. The focus here is on people.

Thus the model could be a temporary project (1); an infrastructure licensed from an aggregator (6); a locally created infrastructure to support projects on valued content (3); or an infrastructure to support content collection and analysis (5); an ongoing integrated infrastructure (7); or something more akin to a community (8).

It could focus on local content (2,3) or published or licensed content (4,6,7) or be content agnostic (5,8). The collection of content could be made by an information service such as a library (2,3,7) by a third party (4,6) or by researchers (1,5). It could focus on temporary (1) or long-term purposes (7).

It could focus on content; technical infrastructure or tools; on skills and collaboration; or purposes. Or it could seek to provide all these.

In reality there might be a mixing of these options, e.g. where some elements of support (8) are supplied alongside a content platform (2,3) or licensing a publisher platform (6).

CASE STUDY

Library research collections and the creation of a community of data scientists

The way that research libraries can support digital humanists and other scholars is a case of how information services can be involved in the application of AI, and one on which much literature has been published (model 3 in table 3.3) (Cordell, 2020; Lewis et al., 2015; Maxwell et al., 2018; Oliver et al., 2019; Padilla, 2019; Wang, 2013). Rich research collections offer many opportunities for the application of computational techniques, and there are many benefits to such projects for libraries in terms of reputation, creating partnerships, improving systems, learning about methods and digitisation, and acquiring new descriptive metadata as a derivative from project outcomes (Ridge, 2019). There are still challenges in terms of copyright and intellectual property rights (IPR), the General Data Protection Regulation (GDPR), in the cost of working at scale, of reintegrating data from projects into systems. Algorithms based on contemporary data do not necessarily work on historic data, so there is a need for creating a significant amount of training data. Skills gaps in library staff can be addressed through running internal courses with a strong hands-on element.

The library or information service could be central to this activity because of its collections, but in some cases there could be other reasons. Fenlon (LIBER, 2020) sees the library as a natural focus for AI partly because of its role in licensing and copyright is important in dealing with publishers of content to be mined, notwithstanding the text and data mining exception in UK law. They can offer an alternative to expensive publisher and aggregator offerings (model 6 in table 3.3). This makes the library central not just for digital humanists using the collection, but to the widening range of disciplines that are using AI. Researchers from many disciplines are starting to use AI, but need training and guidance on tools to use, that libraries could supply. This is analogous to traditional roles in training and advice. The library or information service might also be in a position to fill the technical “infrastructure gap” by providing space to store text data that is being mined (model 5 in table 5.3). This might be better produced at a sector level (model 4 in table 3.3). Finally, as a “neutral” space the library is well-placed to bring together users interested in AI to work together and support collaboration.

With the use of data science techniques across disciplines (but not necessarily rooted in using library special collections or even licensed content) the centre of gravity may have shifted to broader support of data science. Oliver et al. (2019) argue that the academic library is the natural locus for supporting the spread of data science skills across disciplines combined with the promotion of open science principles. This is partly to be achieved by new collaborations within and outside the institution. It illustrates model 8 in table 3.3.

Some of the forces shaping how this service might evolve would be:

- The purposes of analysis
- Cost or type of cost
- Skills base of the information service and the users

3.4 AI interacting with users: new ways to connect

3.4.1 Conversational agents aka chatbots

Conversational agents or chatbots are tools that are programmed to simulate conversation, usually through text interfaces such as instant messaging, or voice. They are widely used on commercial web sites to respond to initial customer queries. They have some obvious advantages in terms of being tireless and available 24/7 (McNeal & Newyear, 2013; Vincze, 2017). They can respond at the point of need, reflecting users’ preference for this style of support. They can pick up routine questions, saving time for more complex queries. For certain purposes the way that the user can remain anonymous may be an advantage. Data from interactions can be mined for further information. They might be delivered through apps or on web sites.

Ethical issues with chatbots include how far to make it explicit that the chatbot is not human and also how to design them not to reinforce gender stereotypes through their naming and behaviour.

It is obvious that there are many potential applications of chatbots in handling user enquires. It is clear too that information professionals might well have an important role to play in building knowledge bases for chatbots. However, despite being by repute in widespread commercial use, they seem to be less used in information contexts, such as libraries. This could be because queries are often too complex or because of a desire to build up relationships with the user. The difficulty of ensuring that answers address the user need accurately or fully might be an issue in some information contexts, e.g. where a high degree of accuracy is required. The effort required to build up a knowledge base of answers or the technical challenges of building them could also be obstacles. There might be particular circumstances in which it is advantageous to build, e.g. in very busy services.

CASE STUDY

The Health Service Executive's LAMA

LAMA (Library Ask Me Anything) is a customised version of Watson Assistant and hosted on the Health Service Executive library website as a chat interface. It has been programmed to respond to a range of library enquiries and to redirect enquirers to the relevant sections of the library website. It has also been configured to recognise symptoms and disease terms and pass a search through to a point of care tool, BMJ Best Practice. It will also recognise common drug terms and search in Medicines Complete. The primary use case for the AI assistant was to ease staff workload, reduce duplication and free staff time for in depth research enquiries and service development. A number of specific targeted use cases were identified:

1. Responding to routine enquiries about OpenAthens accounts (setup, password activation and expiry).
2. Responding to introductory enquiries about library services and access.
3. Responding and direct to further information on locating library resources from specific discovery sources.
4. Responding to the staffed virtual desk chat service where relevant.
5. Enabling AI powered discovery search within chat interface from a list of library resources.
6. Guiding users towards appropriate resources to address their enquiry.

These have all be achieved to some extent and are being built upon.

CASE STUDY

Roche Products Limited: LibraryAnn

LibraryAnn is a chatbot designed by a one-person librarian working in a specialist pharmaceutical company. Their work involves answering copyright enquiries and supporting the UK organisation with finding and accessing medical and developmental resources. Because they are a one-person librarian for a global organisation, they saw the value in an internal chatbot service that could handle routine questions and respond continuously, including outside the librarian's working hours. LibraryAnn answers around three questions a week on chat which represents about half of the weekly information requests. Maintenance takes around an hour a month.

3.4.2 Voice assistants

Voice assistants, also known as smart speakers, such as Amazon's Alexa and Google Home, use automatic speech recognition technology to respond to verbal requests for information. Because they are becoming familiar through the growing level of mass ownership, they have potential to be used by information services. Information services can add their own sets of answers to these tools (Geary, 2019; Shih, 2020). Voice assistants could respond to simple requests, offer wayfinding information or be used to deliver skills tests. Voice search has also been applied to search library and information service catalogues. Voice activated tools could be particularly valuable to certain user groups, such as the elderly. There may be privacy and security issues and concerns arising from their differential ability to understand certain accents.

3.5 AI managing people: questions about ethical use

3.5.1 User management

Information services collect or have potential access to a vast amount of data about their users. This could also be used by AI to help describe, predict or influence behaviour. However, the use of AI to manage user experience is a far more contentious area than analysis of content. While there are ethical, safety and legal challenges to the use of AI in information content, these are much stronger when AI and robots use personal data about users or seek to "nudge" people's behaviour.

Experience in the area of library/ learning analytics in the educational sector may be instructive here. Libraries potentially collect a vast amount of data about university students such as through turnstiles, book circulation, online resource usage (including detailed ebook reading behaviours), interaction with staff or online chat. This could be combined with other institutional data about students, including their demographic data, coursework performance, detailed online learning behaviour and even social media activities. Some researchers have been strong advocates for combining such data to inform decisions about services and to demonstrate library value.

AI could enhance this, e.g. by analysing chat log data, or be used to nudge students towards useful content or to personalise material for users based on their patterns of previous use. In practice, both library and learning analytics seem to be underdeveloped.

There could be a number of explanations of this. There are practical obstacles such as data silos and lack of interoperability (Oakleaf, 2018). Tsai and Gasevic (2017) identify a number of key organisational barriers to the use of learning analytics, including the difficulty institutionally of leading such a profound change management process, the fact that teaching staff are not trained in using them and a lack of evidence for a return on investment in the technologies. However, from a library perspective, what may be critical are the ethical issues:

- The surveillance implied by monitoring library behaviours may inhibit students from searching for particular topics and so freedom of expression.
- The benefits mostly accrue to the university in terms of efficiencies and sales, rather than the students (Jones & Salo, 2018).
- Consent seems rarely to be obtained. Students do not seem to be aware that learning analytics are being used by their teachers.
- Brooke Robertshaw and Asher (2019) argue that the rigour of the data analysis in library analytics studies is suspect and that the benefits in lessons learned do not outweigh the ethical risks.
- Local policies rarely fully reflect ethical concerns (Jones et al., 2020) and studies have rarely been approved by local ethical review boards (Jones, 2019).

Jones comments (2019: 419): "Pursuing socio-technical practices that collect, aggregate, analyse, and act on data revealing students' intellectual behaviors and interests is antithetical to the library profession's commitments to user privacy and intellectual freedom."

This foregrounds the ethical dilemmas around AI, further explored in the next section.

There may be information contexts where such AI applications feels safer, again pointing to the differential impact of AI across information, knowledge management and library sectors. For example, monitoring social media responses to public library services through tools such as sentiment analysis might be deemed a legitimate use, if one regards social media content as in some sense published. Murphy and Villaspesa's (2020) toolkit for museums suggests applications to analyse TripAdvisor experiences or modelling use of museum space (French & Villaespesa, 2019). These may have some application in public library settings. Text mining of survey data is uncontroversial because permission has already been given in consenting to participate in the survey (Moore, 2017). Much of the barrier might be cultural. In the legal sector one use of AI is in more accurately capturing lawyers' time.

3.5.2 Human resource management

Information services such as libraries often have their own staff, and in some cases their own relatively autonomous HR function. Although it is not central to this study, it is worth reflecting on the potential uses of AI in the management of staff, partly because they reinforce the concerns reflected in the previous sub-section. Faggella (2019) identifies some of the possible AI applications in HR, including in assessing applicants, on-boarding, monitoring engagement and predicting attrition. Yet Tambe et al. (2019) point to a number of challenges to applying AI to HR, including the difficulty of quantifying aspects such as human performance, the small amount of data most organisations have on which to train AI (and this may well be deemed to instantiate past biases in recruitment), and the fact that hiring and firing decisions have such a strong effect on human lives, meaning that the ethical issues become acute. While not a central concern for this report, the ethics of AI in HRM is an important dimension of the issues in using AI.

> **Just as the first applications of AI that we have probably experienced are in web and mobile search, so the most obvious application is in information service discovery systems.**

3.5.3 Marketing

The promotion of information services such as libraries is so important that a brief sub-section on the application of AI in marketing was felt to be justified. Ma and Sun (2020) describe a customer's journey towards a purchase, unknowingly shaped by AI: at the stage of an initial search (results of which are partly influenced by company advertising), through customised content displayed on an e-commerce site, by follow up ads and through chatbot interactions. AI enables advertising to be more interactive, personalised and targeted. AI will be used strategically throughout the marketing process (Huang & Rust, 2020). There is potential for the power of such methods to be used by information professionals in promoting services. Equally, they could be perceived to be dimensions of the issue around algorithmic literacy discussed in 3.1.

3.6 Robotic process automation (RPA)

RPA is software that mimics the "path taken by a human through a range of computer applications when performing certain tasks in a business process" (Syed et al., 2020: 1). In that sense it works in a way akin to a macro, except across multiple software applications. A simple example would be to produce a software robot to repeat the process of taking an item from a list, using it to search on a web site, copying the text of the search result and putting it into a spreadsheet or sending an email including it. RPA automates repetitive, routine back office tasks, that have to be repeated many times. It is used for tasks such as data entry and validation, file manipulation, web scraping, data quality checking, etc. An area of particular interest is in automating information governance compliance, e.g. by creating alerts when improper or suspicious activity is detected (AIIM, n.d., 2019). Applications are particularly popular in the financial sector. They ensure accuracy, increase data quality and standardisation, but they are flexible enough to be quickly reconfigured. The software is increasingly easy to use, with Microsoft recently launching Power Automate as part of the Office suite.

Library-specific applications in bibliometrics work are described by Tam (2020). RPA could complement other AI applications to automate parts of workflows (Syed et al., 2020).

RPA relieves the burden of very routine administrative tasks. This could clearly lead to loss of tasks and so perhaps jobs, at least routine ones. It could also lead to job enrichment or free up time for staff to undertake higher level tasks. It clearly creates a role for those who carry out and maintain the automation system; some information professionals might be involved in this work. Schlegel and Kraus (2020) identify the skills required to work on automating tasks.

CASE STUDY

The ad hoc application of AI to a specific information challenge

An interesting example is described in the Government Information Group blog (2020). In this case a vast amount of data was being sent daily at the beginning of the COVID pandemic to the Foreign and Commonwealth Office crisis email from the UK's 270 overseas missions. A team of information professionals used Power Automate to extract data from emails and a wide range of resources into a COVID-19 library which could be accessed by many different stakeholders.

3.7 Smart spaces and robotics

3.7.1 Smart spaces

A smart building uses sensors to collect data about its status and usage and may use this to control lighting and heating, offer services such as wayfinding, or even power interventions that seek to shape user behaviour (Hoy, 2016; Min-Allah & Alrashed, 2020). There are obvious benefits in terms of efficiency savings and sustainability, but potentially also in terms of providing the basis for new information services.

The smart building concept is clearly relevant to information services such as libraries, because in many cases they remain significant physical buildings. Sensors can monitor movement, check occupancy or measure sound levels. This could help understand how users are moving through or settling in a library, with a view to improving its design. It could help predict occupancy, and so help users decide when to come to the library.

It could offer services such as to allow users to book spaces or for wayfinding. It could be used to nudge behaviour, such as to encourage users who have not moved for a long time to take a break, for their well-being. The further one moves into the realms of seeking to influence behaviour, the more smart technologies raise the types of ethical concerns discussed in section 5.5. The smart library might be part of a smart or intelligent campus and ultimately a smart city, with data being integrated across these different levels (JISC, 2019a). Many scenarios of use are imagined in JISC (2019) and the Intelligent Campus blog (intelligentcampus.jiscinvolve.org/wp). Some are as simple as providing integrated travel information. Others are more futuristic.

3.7.2 Robots and libraries

Vlachos et al. (2020) propose that there are three main uses of robots in a library context: for navigation, book location and placement; as information desks; and for learning (see also Tella, 2020; Zheng, 2019). Reflecting on the first usage, we can acknowledge that book returns sorters have existed for some time. Other uses relating to traditional book stock, are to:

1. Retrieve books within large book storage facilities (automated storage retrieval systems)
2. Locate books on standard shelves
3. Scan shelves to identify misplaced and lost books (Liau, 2019)
4. Move large quantities of books (Liau, 2019)

An embodied robot could also be used to welcome or inform users. A telepresence robot enables a remote user to navigate around a space (Guth & Vander Meer, 2017).

An example of a learning use of robots would be to build a robot, perhaps from a kit, as an educational exercise within an information service or makerspace. Robots such as drones can be made available as a borrowable item.

It seems likely that robots developed for very generic tasks in other sectors, rather than robots for library specific tasks, will be the first to reach a scale of production where they become cheap enough for widespread use.

Thus, it is easy to imagine robots being employed in library buildings for:

- Cleaning
- Delivery, including by drones (Nath, 2018)
- Building maintenance
- Robotic furniture (such as furniture that automatically adjusts to the height of the user or shelves that move down to the user)

3.8 4Th industrial revolution automation: the need for AI, robotics and data literacy

AI and robots are being applied in a very wide range of social and organisational settings: from supermarkets to governments.

One important implication for all information professionals is that a basic form of literacy that they need is understanding of AI and robots: they need AI and robotics literacy. Algorithmic literacy could be seen as a subset of this. Given its dependence on big data, AI literacy also implies data literacy. As citizens and employees, information professionals need this understanding themselves. More specifically, AI, robotics and data literacy is something that libraries can play a role in promoting.

CASE STUDY

Create Crates

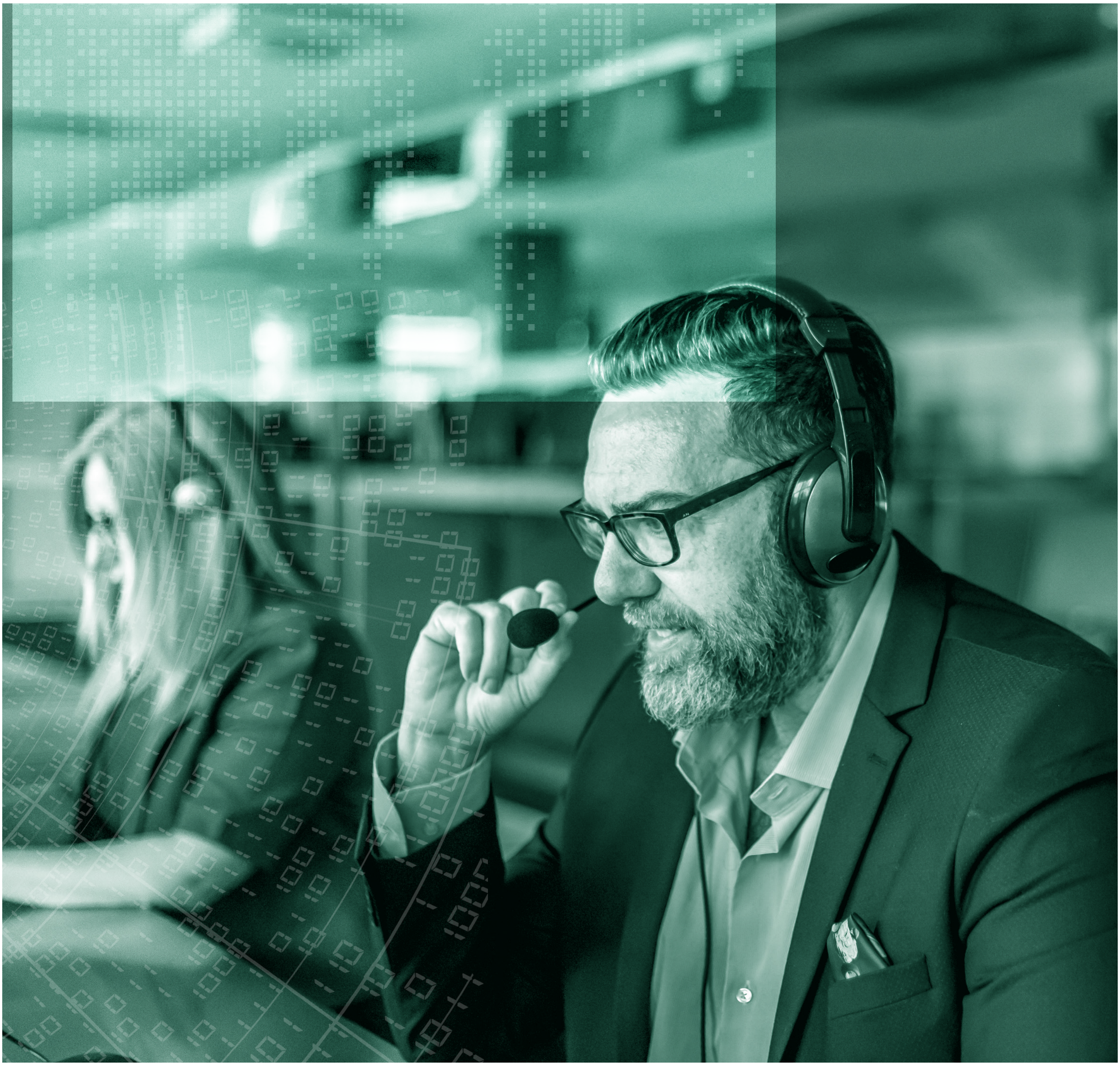
Wakefield public libraries run a service called Create Crates, circulating maker materials to libraries that have neither the space or budget to accommodate permanent makerspaces. Each of the crates is broadly based on a Science, Technology, Engineering or Maths (STEM) subject and allows branches to deliver fun, creative, STEM-based activities that support young people's educational and social development while encouraging innovation, critical thinking, imagination and inventive problem solving without the need for permanently dedicated space or a budget for materials. Each crate includes the core materials necessary to deliver STEM-based activities and ranges from low-tech items such as modelling clay and Lego to high-tech items such as Raspberry Pis, electronic kits and a 3D printer. There are four types of positive impact: firstly, in enthusing children and their parents in STEM, including, potentially, robots and AI; secondl,y in presenting the library service as a place where people are encouraged to create, experiment and explore ideas; thirdly, in boosting library staff skills and confidence in dealing with topics such as AI; and fourthly in creating collaborations with new partners.

This is an obvious role for public libraries since their constituency is the whole population. Makerspaces are the ideal place to promote in-depth exploration, but less intensive introductions would also be relevant. But given the wide and deep impacts of AI and data, all information services could arguably integrate AI and data literacy into their information literacy training.

Appendices 1 and 2 summarise respectively Long and Magerko's (2020) definition of the dimensions of AI literacy (which includes understanding of robotics) and Prado and Marzo's (2013) data literacy definition (others are available).

4

The challenges and opportunities



4 The challenges and opportunities

As the previous section shows, AI and robots have potential to reshape information work in many ways, including some fundamental, even disruptive ones. Yet often the conversations for this report had the mood of the "trough of disillusion" in the Gartner hype cycle. While the technologies hold great future promise, there are recognised to be many challenges to realising these soon; some technical, some organisational and some ethical. Although we are talking about a range of different applications of AI in some very different contexts and in which information professionals are positioned differently, there seems to be some commonality in what these challenges are.

These challenges can also be seen as opportunities for information and knowledge professionals to step forward and contribute or even lead within their context.

4.1 The hype challenge: the need for understanding and a vision

There remains scepticism that many of the products being labelled as AI are truly novel or can fully deliver on vendors' promises. They are often perceived to be familiar technologies rebadged. If they do offer something novel it is more limited than the claim. How proprietary systems work is often a secret. As Callister (2020: 209) comments in the legal context:

"The best we librarians can do in the face of uncertainty [about what approaches to AI vendors are actually using] is to teach our users about the limitations of these systems, disillusioning them of computer intelligence doing the work for them—at least for now. If anything, AI is a tool and, one day perhaps—assuming a humanistic techno-central vision—a partner."

Information professionals often act as a bridge or translator between different groups, such as IT services and domain experts.

There is also a lack of proven evidence of benefit: do they actually do something users or organisations need? Information professionals are well positioned to understand what is really needed by users. There are still questions around where the technology can make a return on investment, be that through business transformation or simply through efficiencies. Yet sometimes the decision makers who invest in AI do not see through the hype and are attracted by the "shiny tech". This can create the pattern of a technical solutionism: technologies in search of a problem.

In some other sectors, such as the academic sector, there seems to be some interest across all disciplines in applications such as machine learning to analyse published texts or other data collections, but it seems to remain unclear which of the computational techniques are relevant (Maxwell et al., 2018).

In other contexts it is hard to find attention for such developments, when there are many competing priorities (Carter, 2018).

The hype challenge creates an opportunity for information professionals to play a role in developing public and organisational understanding of AI and data, to identify where the real benefits of AI lie and so have an informed vision for how they can be implemented in ways aligned to societal need and organisational mission and strategy.

The lack of vision is not universal. One exception to this picture seems to be the health sector, where the NHS is seeing its future as very much digital.

4.2 The ownership challenge: the need for collaboration

Potentially linked to an uncertain vision, it remains unclear where AI sits within the organisation. It might be deemed to be best located in IT. But it could be led by a KM team, by business development or operations, or perhaps it is a strategic project. Some interviewees thought AI sat purely with specialist data scientists.

This creates an opportunity both because of what information professionals can offer from their own capabilities in any collaborations, but also because they often act as a bridge or translator between different groups, such as IT services and domain experts (Henke et al., 2018).

In the academic sector, there was a sense of a widening circle of researchers interested in using AI. This is likely to grow as more AI applications are applied in the economy and it needs to be taught to students not just used in research. However, there is a lack of coordinated support to do this. There was perceived to be a role in bringing academics from different departments together, as the library remains perceived as a neutral space.

In this respect specialisms like KM, with its focus on building collaborations and sharing knowledge, are likely to be very much needed.

The ownership challenge creates an opportunity for information professionals to play a role as bridges and translators, spanning between the technologists and organisational purposes (Henke et al., 2018; Open Data Institute, 2020b).

4.3 The procurement challenge: the need for expertise in licensing, copyright and IPR

AI solutions (and data) often carry a significant premium (Carter, 2019).

Cost is certainly an issue in the public sector. The vendor solutions are perceived to be very expensive and there is an interest in open source and do-it-yourself solutions, though of course these have different types of resource implications. Data processing and infrastructure costs of AI could be high.

Because it is a new market, there is little understanding of the criteria to apply in procurement (Carter, 2019). The players are many and unfamiliar.

In the academic sector there is a proliferation of free and open source tools, but it is hard to navigate to which are the best for particular tasks.

In fact, the procurement could also be of data as much as algorithms.

Although UK law has an exception for text and data mining for non-commercial research, there remain complexities around the law, e.g. if working in a collaboration, how far partners can use licensed content (Caspers et al., 2015; Kelly, 2016; LIBER, 2020; McNeice, 2017b, 2017a). Publishers and aggregators have been accused of using techniques to block or inhibit the legal right to mine, but the law is sufficiently ambiguous to allow them to do this.

The procurement challenge creates an opportunity for information professionals to play a role in procuring both systems and data, drawing on their expertise in procuring systems, licensing content, and their understanding of copyright and IPR.

4.4 The data challenges: the need for data stewardship

AI requires data. The internet giants such as Google and Amazon have vast quantities of data to train their algorithms. This is where AI is already being applied successfully. But within individual organisations the amount of data to feed AI is much less, perhaps not enough to produce effective results. This may place a limit on what is possible. In the commercial sector organisations are unlikely to collaborate and share data. Interestingly, one respondent to Carter (2019: 101) talked about the way that GDPR was "making big data disappear". Even in the public sector there also remain challenges around the openness of data.

In all sectors data that is available is often fragmented and poorly described and organised. But for it to be used by AI it does need to be managed better. This implies a need for organisations to improve their data management in order to leverage AI.

In the academic sector, AI could be applied to historic special collections, but these are very heterogenous (Cordell, 2020).

Within individual organisations the amount of data to feed AI is perhaps not enough to produce effective results. This may place a limit on what is possible.

AI has only been adopted in a few specific areas at any scale. The unexplainability and "opacity" of AI is problematic for its acceptance and use.

The algorithms for image recognition trained on contemporary images do not necessarily work for historical images. As a result, off-the-shelf algorithms cannot be easily applied. Training data is needed and in fact, in all contexts, new types of analysis on particular genre of text will require a lengthy process of training.

In the research library context the provenance of the collections being used as data has to be understood for the results of analysis through AI to be understandable. "Biases" in how historic collections were built up and also in what has been chosen to be digitised in the past need to be taken into account in interpreting analysis.

AI also creates derivative data. For example, in a research library context, AI use on digitised collections creates new descriptive data. Reintegrating this into discovery systems poses issues of quality.

The data challenges create an opportunity for information professionals in terms of data governance and management and long term preservation that might be summarised in the term 'data stewardship'. This builds on existing expertise in the information profession in information governance and information description (taxonomy building, use of standards). Since the data is often information content, collection skills are also highly relevant.

4.5 The technical infrastructure and workflow challenges: the need for information architecting

In many industries current processes are complex and have strong local adaptations, involve lots of human judgement and remain somewhat undocumented. They have not been looked at with an eye to routinise them. In the commercial sector, offshored activities may be more routinised and so more open to AI applications, such as RPA. There is much to be learned about the implementation challenges (Pelz-Shapre & Kompella, 2019).

In the research library context, building workflows for AI systems is also a challenge. It may be that libraries themselves have to create an infrastructure for text and data mining. In the special collection context, it is also recognised that there is a value in integrating crowdsourcing, e.g. to help train AI. But this implies understanding how to motivate potential participants. So it is far from being purely an IT infrastructure that is needed.

The workflow and infrastructure challenges imply the need for information architects.

4.6 The post-implementation challenges: the need for support, training and promotion

Gaining acceptance for AI remains an issue to be tackled. AI has only been adopted in a few specific areas at any scale.

In many sectors, there are strong cultural barriers to adopting AI because many professional groups feel threatened by it and motivated to be doubtful that it can deliver promises to automate tasks that have always before been seen as requiring human skills to undertake.

The unexplainability and "opacity" of AI is problematic for its acceptance and use (Burrell, 2016). If it is hard to understand how a result was produced, there will be scepticism about how far to trust it. Many information uses require replication: this can be hard where machine learning is involved. In other contexts, there may be a temptation to rely too heavily on the outputs of AI without real understanding of how it works and its limitations (Baker, 2018). Critical information literacy remains relevant.

There are many training needs (Carter, 2018). Domain experts need basic training in how to use AI. AI is not advanced enough to work on its own: there remains a need for there to be a human in the loop. There is also a need for trusted intermediaries who can interpret results. In the academic sector, while there appears to be widening interest in using AI across all disciplines, the skills are lacking to do the analysis and interpret the results.

So, there are emerging needs for support, training and the promotion of AI. These post-implementation challenges create an opportunity for information professionals to facilitate the development of AI and its implementation, e.g. through support, training and promotion. This builds on existing roles of the same sort in supporting information use and promoting information services.

4.7 The ethical use challenge: the need for responsible use rooted in professional values and ethics

Previous sections (2.2 and 3.5) surfaced the cherished societal and professional ethical values that could be endangered by misuse of AI and robots: such as fairness and non-discrimination, privacy, the right to freedom of thought and freedom of expression, explainability and transparency, human control over technology, professional responsibility and the promotion of human values.

Information professionals play a vital role when they reflect on how these issues play out in the context where they are working. Their understanding of the nature of data puts them in a strong position to evaluate the less obvious risks, such as created by datafication.

Privacy is not easily protected where masses of data is being joined up, even if it has been anonymised. Where data is collected about individuals there are risks of creating an atmosphere of surveillance.

There is a great potential for bias in training data which can lead to algorithms simply repeating past errors. The way that historic collections have been created reflected past social values which may now make them problematic, even deeply problematic because they privilege certain voices and reinforce the marginalisation of others (Cordell, 2020; Padilla, 2019). Even digitisation initiatives in retrospect may be deemed to have reflected problematic biases.

In many ways the issues are addressed by the values and ethical principles articulated by CILIP (2018) such as around privacy and avoidance of bias. The development of information skills and information literacy resonate with the need for explainable AI and to ensure freedom of thought and expression. The profession has an important role to play in foregrounding ethical issues and finding ethical and safe ways to use AI and robots for the benefit of society. Yet one could argue that more work needs to be done to explore how the profession's enduring values and principles apply to the challenges of AI, particularly in terms of accountability, explainability and transparency.

Information professionals' understanding of the nature of data puts them in a strong position to evaluate the less obvious risks, such as created by datafication.

5 Competencies needed to take the opportunities presented by AI and robots

5 Competencies needed to take the opportunities presented by AI and robots

Chapter 3 identified a number of specific applications of AI and robots of relevance to information professionals. AI in knowledge discovery is particularly important. Section 6 identified the challenges and opportunities for information professionals in enabling organisations to benefit from new technologies, particularly AI. This chapter considers the implications for skills. It is particularly closely tied to the opportunity presented by AI in knowledge discovery.

5.1 The relevance of existing information professional competencies

Table 5.1 identifies the main activities associated with AI and robots, particularly AI in knowledge discovery, identifies the impacts information professionals can have, and maps them to the PKSB and longstanding areas of information work, building on Cox et al. (2019) and Carter (2020).

TABLE 5.1 The relevance of existing information professional competencies

Area of activity	Impact	PKSB	Role extended*	Challenges
Understanding and vision				
Increasing understanding of AI, robots and data.	Organisations maximise the effective and safe use of AI and robots.	Teaching and training skills, supporting users (Literacies and learning).	Information literacy and digital skills training.	Lack of current understanding amongst users.
Leadership – having a vision for how the information service/ organisation / society can best use AI and be involved in wider AI use.	The organisation positions itself to benefit strategically from the technologies, it identifies where the technology can have an impact in either transforming the business or simply through improved efficiency.	Leadership skills, strategic thinking and evaluation (Leadership and advocacy), Business planning (Strategy, planning and management) and communicating with stakeholders (Customer focus, service design and marketing).	Leadership.	Hard to predict what the technologies will really do and in complex environment where main benefit will be; need a clear understanding of the potential.
Collaboration				
Partnership creation for multi-disciplinary teams of domain experts, IT and other services to support AI.	Rich and supportive collaborations come together to reap the benefits of the technologies.	Partnership development (Leadership and advocacy) and communicating with stakeholders (Customer focus, service design and marketing).	Collaboration.	Differing expectations; internal politics.

*Traditional information role being extended.

Area of activity	Impact	PKSB	Role extended*	Challenges
Procurement				
Copyright – understanding what is legally allowed.	The issues around IPR are managed.	Information rights (Information governance and compliance).	Copyright role.	The law may lag behind technologies.
Licensing of content/ data – negotiating/ checking licences and publicising licence terms.	Appropriate licenses are negotiated and their terms are understood by users.	Copyright, intellectual property and licensing (Information governance and compliance).	Licensing of content.	New layer of concerns in examining licences.
Locating content.	Necessary content is collected.	Collection management (Collection management and development).	Collection building.	Anticipates what researchers need – unlikely? Content contains multiple layers of bias.
Licensing of tools or algorithms.	The process of choosing and procuring tools successfully identifies effective and safe technologies.	Contract management (Strategy, planning and management).	Procurement of systems.	Novel types of system (or data?) require new understanding; wide range of tools available; complex issues in vendor relations (Carter, 2019).
Data stewardship				
Description of provenance of data sources and derivatives.	The nature and "biases" in data collections are understood and considered within analysis.	Cataloguing and resource description; ontologies; metadata (Organising knowledge and information).	Creating and managing metadata.	
Data governance.	The risk and benefit from data are effectively managed.	Information governance and compliance.	Information governance.	Ensuring compliance.
Curation and preservation of content created by AI such as derivatives.	Long term benefits are derived from data produced by AI.	Digitisation, curation and preservation (Records management and archiving).	Digital preservation.	Technological uncertainty; lack of a digital preservation culture.
Data analysis, including visualisation.	Data is used for useful analytic descriptions and predictions.	Data analytics (Using and exploiting knowledge and information).		Domain experts, data scientists and AI experts more likely to lead on this but some level of understanding useful.

*Traditional information role being extended.

Area of activity	Impact	PKSB	Role extended*	Challenges
Technical infrastructure				
Infrastructure building/ procurement, e.g. storage for scale and heterogeneity of data.	Effective, reliable, secure and safe technologies are assembled for AI to work within.	Information architecture (Organising knowledge and information), information retrieval (Using and exploiting knowledge and information), data management (Knowledge and information management), ICT skills (IT and communication).	Management of information content specific infrastructure.	IT team more obvious to manage.
Workflow design, including reintegration of derived data into discovery systems.	Workable workflows are created.	Information architecture (Organising knowledge and information), data management (Knowledge and information management), ICT skills (IT and communication).	Procurement of systems.	IT team more obvious to manage.
Support, training and promotion				
Skills development for users.	Users are supported to acquire the skills they need to manage data and use AI tools.	Teaching and training skills, supporting users (Literacies and learning).	Training.	Diverse ways machine learning used makes this challenging. Users such as researchers may not expect to come to the library for this training.
Curation and preservation of content created by AI such as derivatives.	Long term benefits are derived from data produced by AI.	Digitisation, curation and preservation (Records management and archiving)	Digital preservation.	Technological uncertainty; lack of a digital preservation culture.
Marketing new tools to users.	Users are aware of the changing technology options available.	Strategic marketing (Customer focus, service design and marketing).	Marketing of new information services and products.	Finding bandwidth.
Responsible use				
Ethics and values – ensuring that AI uses are ethical and in tune with human and professional values.	Professional values and ethics are respected in every aspect of technology use.	Access to knowledge, intellectual freedom, information skills and information literacy Research ethics (Research skills).	Access to knowledge.	The way familiar issues are manifested may be shifting.

*Traditional information role being extended.

As section 3.3 established, what AI will look like in different contexts could be quite different. The application could be very specific to information service work, e.g. use of AI on a special collection in a research library or an intranet in a law firm or data warehouse run by the information service. But it could also relate to wider organisational use of AI in knowledge work with impacts on information work. This work would be led by knowledge engineers with their own, mostly technical skillset (Anton et al., 2020). In the first case the information service will need to have the vision and make the decisions. In other cases this may be happening elsewhere, but that is all the more reason for the information service to be involved in the project. It may be an individual information professional seeking to get involved in an intelligent automation project.

In many cases this work feels like an extension of some familiar information professional roles, e.g. work around collections, copyright and metadata management. As it is likely to be a collaborative effort there is an emphasis on multi-professional working. The final column surfaces issues such as how it is problematic to convert traditional competencies to work in this new area, including because other professional groups have more obvious expertise.

5.2 Computational thinking, fusion skills and soft skills

Many participants in the study thought that it would be useful if information professionals "got their hands dirty" with AI tools, exploring how to use them to better understand potential applications. AI tools such as for NLP and computer vision have become much easier to use out of the box in the last five years, so this is now possible. Some thought a deeper understanding was required. A classic articulation of this is the notion of "computational thinking", which seeks to capture at an abstract level what type of thinking is needed to solve problems using computers (Barr et al., 2011; Wing, 2006). Another useful conceptualisation of how information professionals might position themselves in relation to computing is "computational sense" (Twidale & Nichols, 2006).

This consists of:

- Comfort and fluency with computational systems
- Metacognitive skills in learning about new computational resources
- Fluency in incremental tailoring and combining of applications for evolving needs
- A sense of applications as ongoing co-designed artefacts rather than technological givens
- A sense of the feasibility of potential design options

This articulates a positioning of an information professional who is comfortable with technologies, learns to use them quickly and plays roles in envisioning their future use, commissions systems, customises systems for local needs and generally bridges between users' needs, organisational purposes, and the perspectives of developers. For many information professionals computational sense may be useful, depending on how technical their role is. This goes beyond the orientation towards IT skills described in the current PKSB, which focuses on specific information service technologies and skills characteristic of a power user. It also aligns to JISC's useful digital capabilities model (JISC, 2019b). Certainly, how the profession positions itself in relation to computing is critical, though it was equally a common opinion among the experts consulted for the report that there was no likelihood or need for information professionals to become computer scientists.

Another perspective is offered by Daugherty and Wilson (2018). Given the ongoing requirement for AI and robots to work together with humans, they see a need for eight "fusion skills" to allow organisations to permit computers and people to work together more effectively (see Appendix 3: Fusion skills). Their last skill of "relentlessly reimagining" is interesting in placing emphasis on the need to reinvent processes to reflect transformed processes, not simply to automate old ones.

While computational thinking or sense, data science skills and even fusion skills may become more important due to AI and robots, there is also a plausible case to say that the skills that will gain in value will be those which computers remain poor at. WEF (Gray, 2016; World Economic Forum, 2020) suggest that the following are areas needed:

- Complex problem solving
- Critical thinking
- Creativity
- People management
- Coordinating with others
- Emotional intelligence
- Judgement and decision making
- Service orientation
- Negotiation
- Cognitive flexibility

This listing reflects the ability to apply intelligence in a broad way (in contrast to the narrow AI achieved by computers) and also to have the soft, human skills that computers are unlikely to develop in the near future. In this context the skills many information professionals possess in collaboration, influencing and negotiating seem to be very valuable. Arguably the PKSB does not currently fully represent some of these higher order cognitive skills; for example, complex problem solving, critical thinking and creativity are not presently identified as professional skills.

The information profession is so diverse and positioned so differently that it is hard to generalise about what skills will be needed, but computational sense, data science skills, fusion skills and these soft skills suggest directions of travel.

5.3 Data science and data stewardship

Use of AI tools to support data analysis has created a new form of analyst role, the data scientist. Data science uses computational methods (such as those discussed in 4.1 above) to derive new knowledge from data. Data scientists tend to have a combination of statistical skills, computational skills and domain knowledge. Some might sit more at the analysis end, some more at the end of business need, and some bridge or translate between the two (Henke et al., 2018; Open Data Institute, 2020a). Data visualisation is also part of what they do. This might be an area where information professionals need to reposition their skills base. Some understanding of the principles involved and hands-on experience of the many tools available will be useful, be that to support users or to apply them directly to their own professional work.

In terms of exploiting data science methods for information service work, there will be strong competition for people with these competencies, so it may be hard to recruit or even retain talent in this area (Markow et al., 2017).

The Edison project (Demchenko et al., 2017) explored skillsets in the broad data science area and drew a useful distinction between data analytics (see Appendix 4: Data Science Data Analysis skills), data management skills (Appendix 5: Data Science Data Management skills) and a number of other skillsets, such as engineering for data science. In reality, as they recognize, roles might sit across these divides: data science is a spectrum (Burton et al., 2017). Information professionals need to be more “data savvy” in general and some information professionals will acquire data analytics skills. However, it seems reasonable to position most information professionals as most likely to work more in the data management than data analysis area, as defined by the Edison project.

As has already been established, the current wave of excitement around AI was preceded by a wave of interest in “big data”, because AI requires large volumes of data. As has also been explained, the range of what is called data has expanded. In AI there is often training data, input data and derived data to manage. As a result, a core aspect of most AI applications is data governance. Since data is closely related to information, data governance and management feels very much within the scope of information professionals. They can certainly contribute a unique perspective in the context of multi-stakeholder data governance (Wendehorst, 2020). Articulating this clearly is important to taking the opportunities presented by AI. Developments in some information sectors, such as around research data management in the academic sector, give us some sense of what this might look like. A lot here may turn on how involved information professionals are already in the data management in their organisation. In some cases the data is library collections or know-how managed by an information service. In other cases there may be other stakeholders who currently manage data.

This recognition should be qualified by remembering that there are other professional groups who would claim the data governance realm, notably computer scientists, IT staff, data analysts and data scientists. One organisation of importance in this context is DAMA. Appendix 6 quotes an overview of DAMA's Data Management Body of Knowledge (DMBoK2) (DAMA International, 2012). The emphasis in DAMA DMBok2 is on the low-level management of data for immediate use, remaining rather detached from what the data is about or its provenance. In a number of ways information professionals might have a rather different view on data governance that might include:

- Emphasis on the provenance of data as created in a particular context and needing to be used in ways based on an understanding of it having originated in that particular context.
- Linked to this, a critical understanding of the limits of any data set: its "bias". The term "data" perhaps implies its "truthiness" that it is "a single, direct, objective representation of a measurable reality" (Fiore-Gartland & Neff, 2015: 476). Actually, this is not true of any form of data, but we particularly know that texts do not have this quality. Individual texts are "biased" by who authored them, in the sense that they obviously represent views of an author or authors. Collections also inevitably have their "biases" that need to be understood in interpreting them (Cordell, 2020; Padilla, 2019). There is a history of how they were created and selected, but when texts become seen as data, this sense that bias needs to be explained and managed may be lost – just as there may be a danger of forgetting the bias, the inherent limits, of the algorithms used by AI to analyse them (Mordell, 2019).
- Emphasis on subject description of the content of data, on categorisation, taxonomies, etc. Computers are increasingly able to do this, but in many contexts building on human concepts.
- An emphasis in many, though not all contexts, on the value of sharing data, even of open data, including based on FAIR principles. The value of commitment to standards, unique identifiers and interoperability is highly relevant.
- Ethics might be given more emphasis than security.
- An emphasis on a lifecycle of data, including a concern with long term preservation.

Thus, information professionals' approach to data is somewhat different from that of DAMA. This kind of appreciation is better articulated in the Edison project's (Demchenko et al., 2017) definition of Data Science Data Management because it mentions issues such as ethics and data curation (Appendix 5). Several participants for this report recognised that librarians have an instinctive understanding of issues relating to the nature of data. To capture the unique information professional perspective on these matters the use of the term 'data stewardship' is suggested. Currently the PKSB locates data management under Knowledge and Information Management, but states that by data we mean structured, usually numerical data. This needs to be extended to reflect the societal re-evaluation of data and to articulate the information profession's unique perspective of data stewardship.

The data librarianship literature explores the way that one branch of the information profession has adapted to work in the area of data management and curation, chiefly for research data, hence the other term used to define work in this area, Research Data Management (Cox, Kennan, et al., 2019; Federer, 2018; Semeler et al., 2019; Tammaro et al., 2019). As well as capturing the technical skills required, the literature also addresses the generic skills and personal characteristics needed by the professional working in this area.

6 Strengths and vulnerabilities of the profession

6 Strengths and vulnerabilities of the profession

The fundamental strength of information professionals in taking up the opportunities created by AI is that, as the previous chapter above indicates, they align to roles they already play. The more similar to traditional roles new activities are, the more probable the competencies already exist; and the more likely that other stakeholders immediately understand that there is a role for information professionals.

6.1 Strengths

Reinforcing this fundamental point we can reflect that:

1. The profession's work in understanding how to navigate across complex ecologies of information and knowledge production and to critically evaluate content is all the more important in a data and information rich world.

2. Web and mobile search has been one of the first areas to see applications of AI. Thus, information professionals have a profound, critical grasp of the underlying issues from promoting user understanding of the value and limits of search engines as an aspect of information literacy. The same kind of issues are apparent in other AI applications.

3. Much of the latest development in AI is around creating, using and especially describing textual content, as well as other forms of content, so closely relates to information professionals' work.

4. Because data is critical to AI, our general understanding of information governance and management more generally is highly relevant, albeit sometimes it has to be translated to be applied to a new context. We have a broad understanding of this as data stewardship.

5. Because there is a strong emphasis on information professional work as a service for users aligned to organisational purposes, we are in a good position to contribute to identifying useful applications of AI, though this might look very different in different contexts.

6. Many of the soft skills that information professionals excel at are all the more relevant because computers are not likely to replicate them in the near future.

7. In terms of adapting to the new technologies, in many sectors information professionals are active in sharing good practice across different organisations, so are able to learn from each other, collectively. The shared professional identity among information, knowledge management and library professionals is an asset, whereas other groups, such as data scientists, lack this coherent professional community.

Engagement with AI and robot applications reflects a desire by information professionals to ensure that we remain seen as a modern, future-facing and digital profession.

> **The strength of information professionals in taking up the opportunities created by AI is that they align to roles they already play.**

6.1 Vulnerabilities

Nevertheless, there are areas where the profession can be seen as vulnerable:

1. While aligned to many roles which information professionals already play, there is a need to translate relevant skills to the new context of AI. This repositioning requires effort and experiment on behalf of organisations and individuals. Organisations in which information professionals are embedded will not always see how they are relevant, especially if they work in something called a library. This returns us to the longstanding dilemma about the naming of the profession: there is respect for the label of librarian, but little public understanding of what it means. There is a collective need to assert the value and relevance of the work of information and knowledge management and library professionals. It is a call often heard, but is nevertheless true.

2. As chapter 5 showed, there is a strong relation between roles in AI and existing aspects of the PKSB. Yet there were clearly areas where the PKSB could be strengthened to reflect a repositioning of the profession.

3. As noted above (in section 2.2, 3.5, 4.7) professional values and ethics are highly relevant to the use of AI but may need some re-articulation to address the challenges AI creates, particularly around accountability, explainability and transparency.
> See recommendation 1.

4. The profession's positioning towards technology may also need to shift. Many participants saw a need for information professionals to get their "hands dirty" trying out the technologies so that they could develop a feel for their potential applications for users. Computational thinking or computational sense is not built into the heart of the profession.
> See recommendation 7.

5. Linked to this, data analysis skills, especially those using computational methods, in data science are at a premium; it will be hard for information services to recruit in these areas.

6. At the same time it is unlikely for most information professionals to get involved in writing algorithms, writing code or doing heavy duty data analysis. So while there is a pull in the direction of embedding computational sense and data science as core professional competencies, there is an equal pull towards the development of the fusion skills to work alongside AI and the soft skills that computers are unlikely to be able to reproduce in the foreseeable future.

7. While information and knowledge professionals have a potential role to play in AI, there are other groups who have a big part to play such as IT specialists, data analysts and data scientists. In organisations where AI is a core competency it is likely that specialist expertise will occupy most of the space around AI. There are also others who claim the data management space (e.g. DAMA). CILIP's constituency is very broad and it is necessary to articulate a vision that resonates in all the different contexts, yet identifies information professionals' unique contribution.
> See recommendation 1.

8. Organisations need to find ways to experiment with AI and technologies that sometimes fail. Some information organisations such as some libraries could be thought of as geared to repeating routine tasks on a massive scale. Creating organisational spaces for innovation within these structures requires different approaches (Cook & Van der Veer Martens, 2019; Rowley, 2011).
> See recommendation 4.

9. Information professionals are busy and resources are stretched. What they already do is important. There are many other new things to be learned, beyond those covered in this report. It has to be understood for which organisations and individuals AI is to be identified as a priority. Individual professionals have to seek out opportunities to try to get to grips with the new technologies in real world applications, and to share their experiences.
> See recommendations 7 & 8.

10. While knowledge sharing and collective learning about professional practice is strong within some sectors of information work, this is not the case in all, and it is weaker between sectors. AI is an agenda across the profession and the conversation could also include adjacent professions such as archivists and museum curators, because in many ways foundational distinctions between types of content are eroding.
> See recommendations 3, 6 & 8.

11. Learning providers, such as library and information schools, who train new professionals need to give them a fundamental grounding in AI, although it is hard to squeeze this into already packed curricula.
> See recommendations 9, 10 & 11.

12. Curricula need to be updated, but the more immediate issue is CPD for current professionals who will be operating over the next decades as AI in its many forms gradually becomes common practice.
> See recommendation 12.

The more immediate issue is CPD for current professionals who will be operating over the next decades as AI in its many forms gradually becomes common practice.

Recommendations

The technologies discussed in this report, though not without their challenges, present a major opportunity for social good, as the profession learns more about how to use them and supports users to engage productively and safely with them.

In responding to the opportunities and challenges there are a number of areas of potential action for stakeholders.

For CILIP

Recommendation 1

It would be productive if CILIP could facilitate a collective articulation of how the profession can contribute to AI, such as through the role in AI and data literacy, through the notion of data stewardship, and the relevance of its values and ethics to this new context. The current review of the PKSB could strengthen its alignments to the needs of AI and robots, especially in relation to computational sense, data science, data stewardship and the soft skills needed to work effectively with AI and robots. Facilitating discussion within the profession about how CILIP's ethics and value statements apply to the specific case of AI and robots would help in strengthening the profession's stance.

Recommendation 2

This would be supported if CILIP could identify pathfinder organisations and individuals who demonstrate how AI and robots can be introduced for the benefit of users and organisations.

Recommendation 3

CILIP with other professional bodies can foster knowledge sharing across the profession, and with adjacent professions, through events and curating relevant open learning resources. CILIP special interest groups offer a locus for developing communities of practice to support learning about the new technologies.

For information services and libraries

Recommendation 4

Leaders in information services and libraries need to create organisational structures within which experiment is possible and within which individual learning is supported and encouraged.

Recommendation 5

Information services and libraries should actively engage with AI to explore the potential benefit to users.

Recommendation 6

Where possible, information services should actively share their knowledge across the wider profession.

For individual information professionals

Recommendation 7

In the spirit of the professional obligation to keep up to date, individuals need to be inquisitive and willing to explore the new technologies.

Recommendation 8

In the spirit of the professional obligation to engage with colleagues, individuals need to share their understanding and promote their vision of the relevance of the profession.

For educational institutions at all levels

Recommendation 9

Training for new entrants to the profession needs to encompass an understanding of AI and how it may be applied in information contexts.

Recommendation 10

There could be a greater stress on developing computational sense, data science, fusion skills and also on the types of soft skills in greater demand because computers cannot carry them out. There could be a greater stress placed on data stewardship.

Recommendation 11

There is a need for hands-on experiences with AI applications in practice.

For other training providers

Recommendation 12

There is a need both for taster courses and more in-depth training in relevant AI applications.

For research

Recommendation 13

More research is needed on the adoption of AI, the organisational structures that support it and on the impact of AI on information, knowledge management and library professional roles.

References

67 bricks. (2017). *Content Enrichment : An essential strategic capability for every publisher*. www.67bricks.com/content-enrichment-an-essential-strategic-capability-for-every-publisher.

AIIM. (n.d.). *Robotic Process Automation (RPA) for Information Management Professionals: Is RPA Right for My Business?* info.aiim.org/is-rpa-right-for-my-business.

AIIM. (2019). *10 Things You Need to Know About Process Services and RPA*. info.aiim.org/10-things-you-need-to-know-about-process-automation-and-rpa.

AIIM Industry Watch. (2018). *Leveraging Deep Learning and Machine Learning Capabilities*. info.aiim.org/leveraging-deep-learning-and-machine-learning-capabilities.

AINow. (2019). *2019 report*. ainowinstitute.org/AI_Now_2019_Report.pdf.

Anderson, C. B., & Craiglow, H. A. (2017). Text mining in business libraries. *Journal of Business and Finance Librarianship*, 22(2), 149–165. doi.org/10.1080/08963568.2017.1285749.

Anton, E., Behne, A., & Teuteberg, F. (2020). The humans behind artificial intelligence – an operationalisation of AI competencies. *Twenty-Eighth European Conference on Information Systems (ECIS2020), June*. aisel.aisnet.org/ecis2020_rp/141.

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

Baker, J. J. (2018). 2018: A legal research odyssey: Artificial intelligence as disruptor. *Law Library Journal*, 110(1), 5–30. doi.org/10.2139/ssrn.2978703.

Bakke, A. (2020). Everyday Googling: Results of an Observational Study and Applications for Teaching Algorithmic Literacy. Computers and Composition, 57, 102577. doi.org/10.1016/j.compcom.2020.102577.

Barr, D., Harrison, J., & Conery, L. (2011). Computational Thinking: A Digital Age Skill for Everyone. *Learning and Leading with Technology*, 38(6), 20–23. eric.ed.gov/?id=EJ918910.

Beller, E., Clark, J., Tsafnat, G., Adams, C., Diehl, H., Lund, H., Ouzzani, M., Thayer, K., Thomas, J., Turner, T., Xia, J., Robinson, K., Glasziou, P., Ahtirschi, O., Christensen, R., Elliott, J., Graziosi, S., Kuiper, J., Moustgaard, R., … Wedel-Heinen, I. (2018). Making progress with the automation of systematic reviews: Principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). Systematic Reviews, 7(1), 1–7. doi.org/10.1186/s13643-018-0740-7.

British Computing Society. (2020). *AI and your job*. doi.org/10.1093/itnow/bwaa063.

Brooke Robertshaw, M., & Asher, A. (2019). Unethical numbers? A meta-analysis of library learning analytics studies. Library Trends, 68(1), 76–101. doi.org/10.1353/lib.2019.0031.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data and Society*, 3(1), 1–12. doi.org/10.1177/2053951715622512.

Burton, M., Lyon, L., Erdmann, C., & Tijerina, B. (2017). *Shifting to Data Savvy: The Future of Data Science In Libraries*. d-scholarship.pitt.edu/33891.

Callister, P. D. (2020). Law, artificial intelligence, and natural language processing: A funny thing happened on the way to my search results. *Law Library Journal*, 112(2), 161–212. ssrn.com/abstract=3712306.

Carter, D. (2018). How real is the impact of artificial intelligence? The business information survey 2018. *Business Information Review*, 35(3), 99–115. doi.org/10.1177/0266382118790150.

Carter, D. (2019). New technologies and new data sources: The business information survey 2019. *Business Information Review*, 36(3), 95–105. doi.org/10.1177/0266382119871848.

Carter, D. (2020). Regulation and ethics in artificial intelligence and machine learning technologies: Where are we now? Who is responsible? Can the information professional play a role? *Business Information Review*, 37(2), 60–68. doi.org/10.1177/0266382120923962.

Caspers, M., Guibault, L., McNeice, K., Piperidis, S., Pouli, K., Eskevich, M., & Gavrilidou, M. (2015). *Reducing barriers and increasing uptake of text and data mining for research Baseline report of policies and barriers of TDM in Europe (extended version of D3 . 3)*. project.futuretdm.eu/wp-content/uploads/2017/05/FutureTDM_D3.3-Baseline-Report-of-Policies-and-Barriers-of-TDM-in-Europe.pdf.

Centre for Data Ethics and Innovation. (2020). *Review into bias in algorithmic decision-making*. assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/938850/Main_Report_-_CDEI_Review_into_Bias_in_Algorithmic_Decision-Making.pdf.

CILIP. (2018). Ethical Framework. www.cilip.org.uk/general/custom.asp?page=ethics.

Cook, M., & Van der Veer Martens, B. (2019). Managing Exploratory Units in Academic Libraries. *Journal of Library Administration*, 59(6), 606–628. doi.org/10.1080/01930826.2019.1626647.

Coombs, C., Hislop, D., Taneva, S. K., & Barnard, S. (2020). The strategic impacts of Intelligent Automation for knowledge and service work: An interdisciplinary review. *Journal of Strategic Information Systems*, July 2017, 101600. doi.org/10.1016/j.jsis.2020.101600.

Cordell, R. (2020). *Machine Learning + Libraries: A report on the state of the field*. labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf?.

Cox, A. M. (2021). Exploring the impact of Artificial Intelligence and robots on Higher Education through literature-based design fictions. *International Journal of Educational Technology in Higher Education*. doi.org/doi.org/10.1186/s41239-020-00237-8.

Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. (2019). Maturing research data services and the transformation of academic libraries. *Journal of Documentation*, 75(6). doi.org/10.1108/JD-12-2018-0211.

Cox, A. M., Pinfield, S., & Rutter, S. (2019). The intelligent library. *Library Hi Tech*, 37(3), 418–435. doi.org/doi.org/10.1108/LHT-08-2018-0105.

DAMA International. (2012). DAMA-DMBOK2 Framework. In *The Data Management Association*. dama.org/sites/default/files/download/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf.

Daugherty, P. R., & Wilson, H. J. (2018). *Human + Machine: Reimagining work in the age of AI*. Harvard Business Review Press.

Demchenko, Y., Belloum, A., & Wiktorski, T. (2017). *EDISON Data Science Framework: Part 1. Data Science Competence Framework (CF-DS) Release 2* (Issue July). edison-project.eu/sites/edison-project.eu/files/filefield_paths/edison_cf-ds-release2-v08_0.pdf.

Eliot, B., Andrews, W., & Hare, J. (2020). *A framework for applying AI in the enterprise*.

Elliott, J. H., Turner, T., Clavisi, O., Thomas, J., Higgins, J. P. T., Mavergames, C., & Gruen, R. L. (2014). Living systematic reviews: An emerging opportunity to narrow the evidence-practice gap. *PLoS Medicine*, 11(2), 1–6. doi.org/10.1371/journal.pmed.1001603.

Equality Task Force. (2020). *Mind the gap: How to fill the equality and AI accountability gap in an automated world*. www.ifow.org/publications/mind-the-gap-the-final-report-of-the-equality-task-force.

Faggella, D. (2019). *Machine learning in Human Resources – applications and trends*. Emerj. emerj.com/ai-sector-overviews/machine-learning-in-human-resources.

Faggella, D. (2020). *AI in Law and Legal Practice – A Comprehensive View of 35 Current Applications*. Emerj. emerj.com/ai-sector-overviews/ai-in-law-legal-practice-current-applications.

Federer, L. (2018). Defining data librarianship: A survey of competencies, skills, and training. *Journal of the Medical Library Association*, 106(3), 294–303. doi.org/10.5195/jmla.2018.306.

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. doi.org/10.1108/LHTN-05-2016-0024.

Fiore-Gartland, B., & Neff, G. (2015). Communication, mediation, and the expectations of data: Data valences across health and wellness communities. *International Journal of Communication*, 9, 1466–1484. ijoc.org/index.php/ijoc/article/view/2830.

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*. cyber.harvard.edu/publication/2020/principled-ai.

Flannery, J. (2020). Using NLP to Generate MARC Summary Fields for Notre Dame 's Catholic Pamphlets. *International Journal of Librarianship*, 5(1), 20–35. doi.org/10.23974/ijol.2020.vol5.1.158.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1–15. doi.org/10.1162/99608f92.8cd550d1.

French, A., & Villaespesa, E. (2019). AI, Visitor Experience, and Museum Operations: A Closer Look at the Possible. *In Humanizing the Digital: Unproceedings from the MCN 2018 Conference* (pp. 101–113).

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. doi.org/10.1016/j.techfore.2016.08.019.

Geary, D. (2019). How to bring AI into your library. *Infotoday*, September, 32–36.

Glanville, J. (2019). Text mining for information specialists. In P. Levay & J. Craven (Eds.), *Systematic Searching : Practical Ideas for Improving Results* (pp. 147–170). Facet. doi.org/10.29085/9781783303755.008.

Glanville, J., & Wood, H. (2018). *Text Mining Opportunities: White Paper*. www.cadth.ca/text-mining-opportunities-white-paper.

Gramatica, R., & Pickering, R. (2017). Start-up story Yewno: an AI-driven path to a knowledge-based future The launch of Yewno. *UKSG Insights*, 30(2), 107–111. insights.uksg.org/articles/10.1629/uksg.369.

Grames, E. M., Stillman, A. N., Tingley, M. W., & Elphick, C. S. (2019). An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, 10(10), 1645–1654. doi.org/10.1111/2041-210X.13268.

Gray, A. (2016). *The 10 skills you need to thrive in the Fourth Industrial Revolution*. World Economic Forum. www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution.

Group, C. G. I. (2020). GIG *AWARD 2020 : Foreign and commonwealth office's knowledge and information team*. cilipgig.blogspot.com/2020/11/gig-award-2020-foreign-and-commonwealth.html.

Guth, L. M., & Vander Meer, P. (2017). Telepresence robotics in an academic library: A study of exposure and adaptation among patrons and employees. *Library Hi Tech*, 35(3), 408–420. doi.org/10.1108/LHT-03-2017-0059.

Henke, N., Levine, J., & McInerney, P. (2018). Analytics translator: The new must-have role. *Harvard Business Review*.

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education. Center for curriculum redesign.

Hoy, M. B. (2016). Smart Buildings: An Introduction to the Library of the Future. *Medical Reference Services Quarterly*, 35(3), 326–331. doi.org/10.1080/02763869.2016.1189787.

Huang, M. H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 4. doi.org/10.1007/s11747-020-00749-9.

IFLA. (2020). *IFLA Statement on Libraries and Artificial Intelligence*. www.ifla.org/files/assets/faife/ifla_statement_on_libraries_and_artificial_intelligence.pdf.

JISC. (2019a). *Guide to the intelligent campus: Using data to make smarter use of your university or college estate*. repository.jisc.ac.uk/6882/1/Intelligent_Campus_Guide.pdf.

JISC. (2019b). *Jisc digital capabilities framework: The six elements defined*. repository.jisc.ac.uk/7278/1/BDCP-DC-Framework-Individual-6E-110319.pdf.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. doi.org/10.1038/s42256-019-0088-2.

Johnson, J. D. (2020). Theorizing Network Bias and Teaching Mêtic Invention in Online Search. *Computers and Composition*, 56 (August 2018), 102573. doi.org/10.1016/j.compcom.2020.102573.

Jones, K. M. L. (2019). "Just because you can doesn't mean you should": Practitioner perceptions of learning analytics ethics. *Portal*, 19(3), 407–428. doi.org/10.1353/pla.2019.0025.

Jones, K. M. L., Briney, K. A., Goben, A., Salo, D., Asher, A., & Perry, M. R. (2020). A comprehensive primer to library learning analytics practices, initiatives, and privacy issues. *College and Research Libraries*, 81(3), 570–591. doi.org/10.5860/crl.81.3.570.

Jones, K. M. L., & Salo, D. (2018). Learning analytics and the academic library: Professional ethics commitments at a crossroads. *College and Research Libraries*, 79(3), 304–323. doi.org/10.5860/crl.79.3.304.

Jonnalagadda, S. R., Goyal, P., & Huffman, M. D. (2015). Automating data extraction in systematic reviews: A systematic review. *Systematic Reviews*, 4(78). doi.org/10.1186/s13643-015-0066-7.

Kay, D. (2018). *Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment*. www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ReportGA73.aspx.

Kelly, J. (2016). THE TEXT AND DATA MINING COPYRIGHT EXCEPTION: BENEFITS AND IMPLICATIONS FOR UK HIGHER EDUCATION. www.jisc.ac.uk/guides/text-and-data-mining-copyright-exception.

Kroski, E. (2020). *Law librarianship in the age of AI*. American Library Association.

Lewis, V., Spiro, L., Wang, X., & Cawthorne, J. E. (2015). Building Expertise to Support Digital Scholarship: A Global Perspective. In *CLIR Publication* (Issue 168). www.clir.org/pubs/reports/pub168/pub168.

Liau, Y. chin. (2019). *Transforming library operation with robotics*. August, 1–5. creativecommons.org/licenses/by/4.0.

LIBER. (2020). *Research Libraries: How You Can Support Text and Data Mining*.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Conference on Human Factors in Computing Systems - Proceedings*, 1–16. doi.org/10.1145/3313831.3376727.

Lowendahl, J.-M., & Calhoun Williams, K. (2018). *5 best practices for artificial intelligence in higher education*.

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing – Connecting computing power to human insights. *International Journal of Research in Marketing*, 37(3), 481–504. doi.org/10.1016/j.ijresmar.2020.04.005.

Markow, W., Braganza, S., Taska, B., Miller, S. M., & Hughes, D. (2017). *The Quant Crunch: How the Demand for Data Science Skills is Disrupting the Job Market*. www.burning-glass.com/research-project/quant-crunch-data-science-job-market.

Maxwell, D., Norton, H., & Wu, J. (2018). The Data Science Opportunity: Crafting a Holistic Strategy. *Journal of Library Administration*, 58(2), 111–127. doi.org/10.1080/01930826.2017.1412704.

Mckinsey Analytics. (2019). Global AI Survey: AI proves its worth, but few scale impact. In *McKinsey &Company* (Issue November). www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact.

McKinsey Global Institute. (2018). Skill Shift: Automation and the Future of the Workforce (Discussion Paper, May 2018). In *McKinsey &Company* (Issue May). www.mckinsey.com/~/media/McKinsey/Featured Insights/Future of Organizations/Skill shift Automation and the future of the workforce/MGI-Skill-Shift-Automation-and-future-of-the-workforce-May-2018.ashx.

McNeal, M., & Newyear, D. (2013). Introducing chatbots in libraries. *Library Technology Reports*, 49(8), 5–10.

McNeice, K. (2017a). Skills and education: The need for strategic thinking. *FutureTDM Symposium*.

McNeice, K. (2017b). Text and Data Mining: The Changing Nature of Research. *FutureTDM Symposium*.

Millard, T., Synnot, A., Elliott, J., & Turner, T. (2018). *Results from the evaluation of the pilot living systematic reviews*. community.cochrane.org/sites/default/files/uploads/inline-files/Transform/201905 LSR_pilot_evaluation_report.pdf.

Min-Allah, N., & Alrashed, S. (2020). Smart campus—A sketch. *Sustainable Cities and Society*, 59. doi.org/10.1016/j.scs.2020.102231.

Moore, M. T. (2017). Constructing a sentiment analysis model for LibQUAL+ comments. *Performance Measurement and Metrics*, 18(1), 78–87. doi.org/10.1108/PMM-07-2016-0031.

Mordell, D. (2019). Critical questions for archives as (big) data. *Archivaria*, 87(5), 140–161. archivaria.ca/index.php/archivaria/article/view/13673.

Murphy, O., & Villaespesa, E. (2020). *The Museums + AI Network*. themuseumsainetwork.files.wordpress.com/2020/02/20190317_museums-and-ai-toolkit_rl_web.pdf.

Nath, F. (2018). Library drone delivery programme: A study. *DESIDOC Journal of Library and Information Technology*, 38(5), 350–354. doi.org/10.14429/djlit.38.5.12892.

Neary, M. A., & Chen, S. X. (2017). Artificial intelligence: Legal research and law librarians. *AALL Spectrum 21*, 5(6), 16–20. lawdigitalcommons.bc.edu/lsfp/1073.

NHSX. (2019). *Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care*. www.nhsx.nhs.uk/media/documents/NHSX_AI_report.pdf.

O'Connor, A. M., Tsafnat, G., Gilbert, S. B., Thayer, K. A., Shemilt, I., Thomas, J., Glasziou, P., & Wolfe, M. S. (2019). Still moving toward automation of the systematic review process: A summary of discussions at the third meeting of the International Collaboration for Automation of Systematic Reviews (ICASR). *Systematic Reviews*, 8(1), 1–5. doi.org/10.1186/s13643-019-0975-y.

Oakleaf, M. (2018). The problems and promise of learning analytics for increasing and demonstrating library value and impact. *Information and Learning Science*, 119(1–2), 16–24. doi.org/10.1108/ILS-08-2017-0080.

OECD. (2019). *Artificial intelligence in society*. www.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77-en.

Oliver, J. C., Kollen, C., Hickson, B., & Rios, F. (2019). Data Science Support at the Academic Library. *Journal of Library Administration*, 59(3), 241–257. doi.org/10.1080/01930826.2019.1583015.

Open Data Institute. (2020a). *Data Skills Framework*. theodi.org/article/data-skills-framework.

Open Data Institute. (2020b). *Getting data right: Perspectives on the UK National Data Strategy* 2020. library1.nida.ac.th/termpaper6/sd/2554/19755.pdf.

Padilla, T. (2019). Responsible operations: Data science, machine learning, and AI in libraries. In *OCLC*. doi.org/10.25333/xk7z-9g97.

Pelz-Shapre, A., & Kompella, K. (2019). *Practical artificial intelligence: An enterprise playbook*. Deep Publishing.

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123–134. doi.org/10.1515/libri-2013-0010.

Ridge, M. (2019). *Operationalising AI at a national library*. Museums + AI Network. www.slideshare.net/miaridge/operationalising-ai-at-a-national-library.

Rowley, J. (2011). Innovation for survival: From cooperation to collaboration. In *Advances in Librarianship* (Vol. 34). Emerald Group Publishing Ltd. doi.org/10.1108/S0065-2830(2011)0000034013.

Sander, I. (2020). Critical big data literacy tools—Engaging citizens and promoting empowered internet usage. *Data & Policy*, 2. doi.org/10.1017/dap.2020.5.

Schlegel, D., & Kraus, P. (2020). Robotic process automation as an emerging career opportunity: An analysis of required qualifications and skills. *13th Annual Conference of the EuroMed Academy of Business*, 1083–1097. publikationen.reutlingen-university.de/frontdoor/deliver/index/docId/2815/file/2815.pdf.

Semeler, A. R., Pinto, A. L., & Rozados, H. B. F. (2019). Data science in data librarianship: Core competencies of a data librarian. *Journal of Librarianship and Information Science*, 51(3), 771–780. doi.org/10.1177/0961000617742465.

Shih, W. (2020). Voice revolution. Library Technology Reports, 56(4), 5–13.

Smith, A. (2016). Big Data technology, evolving knowledge skills and emerging roles. *Legal Information Management*, 16(4), 219–224. doi.org/doi.org/10.1017/S1472669616000499.

Stanfill, M. H., & Marc, D. T. (2019). Health information management: Implications of artificial intelligence on healthcare data and information management. *Yearbook of Medical Informatics*, 28(1), 56–64. doi.org/10.1055/s-0039-1677913.

Stuart, D. (2020). *Data science for information professionals*. Facet Publishing.

Syed, R., Suriadi, S., Adams, M., Bandara, W., Leemans, S. J. J., Ouyang, C., ter Hofstede, A. H. M., van de Weerd, I., Wynn, M. T., & Reijers, H. A. (2020). Robotic Process Automation: Contemporary themes and challenges. *Computers in Industry*, 115, 103162. doi.org/10.1016/j.compind.2019.103162.

Tam, W. (2020). *The rise of the robots: Using robotic process automation in bibliometrics*. tv.theiet.org/?videoid=14133.

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61(4), 15–42. doi.org/10.1177/0008125619867910.

Tammaro, A. M., Matusiak, K. K., Sposito, F. A., & Casarosa, V. (2019). Data curator's roles and responsibilities: An international perspective. *Libri*, 69(2), 89–104. doi.org/10.1515/libri-2018-0090.

Tella, A. (2020). Robots are coming to the libraries: Are librarians ready to accommodate them? *Library Hi Tech News*, 37(8), 13–17. doi.org/10.1108/LHTN-05-2020-0047.

Thelwall, M. (2018). Artificial Intelligence, Automation and Peer Review. In *Jisc*. link.springer.com/article/10.1007/BF02016664.

Tsai, Y. S., & Gasevic, D. (2017). Learning analytics in higher education - Challenges and policies: A review of eight learning analytics policies. *ACM International Conference Proceeding Series*, 233–242. doi.org/10.1145/3027385.3027400.

Twidale, M. B., & Nichols, D. M. (2006). *Computational sense: The role of technology in the education of digital librarians*. researchcommons.waikato.ac.nz/handle/10289/40.

Vincze, J. (2017). Virtual reference librarians (Chatbots). *Library Hi Tech News*, 34(4), 5–8. doi.org/10.1108/LHTN-03-2017-0016.

Vlachos, E., Faber Hansen, A., & Povl Holck, J. (2020). A robot in the library. In M. Rauterberg (Ed.), *Culture and Computing* (pp. 312-). Springer. doi.org/10.1007/978-3-030-50267-6.

Wagstaff, K. L., & Liu, G. Z. (2018). Automated classification to improve the efficiency of weeding library collections. Journal of Academic Librarianship, 44(2), 238–247. doi.org/10.1016/j.acalib.2018.02.001.

Wang, M. (2013). Supporting the research process through expanded library data services. *Program*, 47(3), 282–303. doi.org/10.1108/PROG-04-2012-0010.

Wendehorst, C. (2020). *Framework paper for GPAI work on data governance*.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. doi.org/10.1145/1999747.1999811.

World Economic Forum. (2020). *The 10 skills you need to thrive in the Fourth Industrial Revolution*. www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution.

Zheng, W. (2019). How do library staff view librarian robotics? Librarian staff's ignored humanistic views on the impact and threat of robotics adoption. *IFLA WLIC 2019*, 1–17. library.ifla.org/2751/1/s02-2019-wang-en.pdf.

Appendixes

APPENDIX 1

AI literacy, according to Long and Magerko (2020)

1. Recognizing AI

Distinguish between technological artifacts that use and do not use AI.

2. Understanding Intelligence

Critically analyze and discuss features that make an entity "intelligent", including discussing differences between human, animal, and machine intelligence.

3. Interdisciplinarity

Recognize that there are many ways to think about and develop "intelligent" machines. Identify a variety of technologies that use AI, including technology spanning cognitive systems, robotics, and ML.

4. General vs. Narrow

Distinguish between general and narrow AI.

5. AI's Strengths & Weaknesses

Identify problem types that AI excels at and problems that are more challenging for AI. Use this information to determine when it is appropriate to use AI and when to leverage human skills.

6. Imagine Future AI

Imagine possible future applications of AI and consider the effects of such applications on the world.

7. Representations

Understand what a knowledge representation is and describe some examples of knowledge representations.

8. Decision-Making

Recognize and describe examples of how computers reason and make decisions.

9. Machine Learning Steps

Understand the steps involved in machine learning and the practices and challenges that each step entails.

10. Human Role in AI

Recognize that humans play an important role in programming, choosing models, and fine-tuning AI systems.

11. Data Literacy

Understand basic data literacy concepts such as those outlined in Prado and Marzal (2013).

12. Learning from Data

Recognize that computers often learn from data (including one's own data).

13. Critically Interpreting Data

Understand that data cannot be taken at face-value and requires interpretation. Describe how the training examples provided in an initial dataset can affect the results of an algorithm.

14. Action & Reaction

Understand that some AI systems have the ability to physically act on the world. This action can be directed by higher-level reasoning (e.g. walking along a planned path) or it can be reactive (e.g. jumping backwards to avoid a sensed obstacle).

15. Sensors

Understand what sensors are, recognize that computers perceive the world using sensors, and identify sensors on a variety of devices. Recognize that different sensors support different types of representation and reasoning about the world.

16. Ethics

Identify and describe different perspectives on the key ethical issues surrounding AI (i.e. privacy, employment,

misinformation, the singularity, ethical decision making, diversity, bias, transparency, accountability).

17. Programmability

Understand that agents are programmable.

Note that Data literacy is explicitly mentioned and Prado and Marzal (2013) cited as a source for this subset of AI literacy.

APPENDIX 2

Data literacy, as defined by Prado and Marzal (2013)

1. Understanding data

What is data?

Competency: learners need to know what is meant by data and be aware of the various possible types of data.

Contents: Data definition; Types of data (depending on origin, format, usage license and so on).

Data in society

Competency: learners need to be aware of the role of data in society, how they are generated and by whom, and their possible applications, as well as the implications of their use.

Contents: Data producers and consumers; Data lifecycle; Data applications: their impact on science and society; Copyright and licenses influencing data reuse.

2. Finding and/or obtaining data

Data sources

Competency: learners need to be aware of the possible data sources, be able to evaluate them and select the ones most relevant to an informational need or a given problem.

Contents: Data sources; Criteria for assessing data sources.

Obtaining data

Competency: learners need to be able to detect when a given problem or need cannot be (totally or partially) solved with the existing data and, as appropriate, undertake research to obtain new data.

Contents: Main research methods for obtaining original data.

3. Reading, interpreting and evaluating data

Reading and interpreting data

Competency: learners need to be aware of the various forms in which data can be presented (written, numerical or graphic), and their respective conventions, and be able to interpret them.

Contents: Ways to present and represent data.

Evaluating data

Competency: learners need to be able to evaluate data critically

Contents: Data evaluation criteria (including authorship, method of obtaining and analyzing data, comparability, inference and data summaries).

4. Managing data

Data and metadata collection and management

Competency: learners need to be aware of the need to save the data selected or generated and of descriptive or other data associated therewith, for due identification, management and subsequent reuse.

Contents: Metadata; Reference management tools; Databases; Data management repositories: policies and practices.

5. Using data

Data handling

Competency: learners need to be able to prepare data for analysis, analyze them in keeping with the results sought and know how to use the necessary tools.

Contents: Data conversion; Handling data analysis tools, both locally (Excel, R, SPSS, Stata or similar) and on-line.

Producing elements for data synthesis

Competency: learners need to be able to synthesize and represent the results of data analysis in ways suited to the nature of the data, their purpose and the audience targeted in the inquiry.

Contents: Choosing suitable data representation methods (tables, graphs or similar); Handling tools (built into analytical tools or stand-alone applications such as Gapminder, Visual.ly or IBM's Many Eyes).

Ethical use of data

Competency: learners need to make ethical use of data, acknowledging the source when obtained or formulated by others, and making sure that used methods are deployed and results interpreted transparently and honestly.

Contents: What is the ethical use of data; How to cite data sources.

APPENDIX 3

Fusion Skills (Daugherty & Wilson, 2018)

Rehumanizing time

The ability to increase the time available for distinctly human tasks like interpersonal interactions, creativity, and decision making in a reimagined business process.

Responsible normalizing

The act of responsibly shaping the purpose and perception of human-machine interaction as it relates to individuals, businesses, and society.

Judgment integration

The judgment-based ability to decide a course of action when a machine is uncertain about what to do.

Intelligent interrogation

Knowing how best to ask questions of AI, across levels of abstraction, to get the insights we need.

Bot-based empowerment

Working well with AI agents to extend your capabilities, and create superpowers in business processes and professional careers.

Holistic melding

The ability to develop robust mental models of AI agents to improve business process outcomes.

Reciprocal apprenticing

1. Performing tasks alongside AI agents so they can learn new skills;

2. On-the-job training for people so they can work well within AI-enhanced processes.

Relentless reimagining

The rigorous discipline of creating new processes and business models from scratch, rather than simply automating old processes.

APPENDIX 4

Edison project (Demchenko et al., 2017): Data science data analytics

DSDA01 Effectively use variety of data analytics techniques, such as Machine Learning (including supervised, unsupervised, semisupervised learning), Data Mining, Prescriptive and Predictive Analytics, for complex data analysis through the whole data lifecycle.

DSDA02 Apply designated quantitative techniques, including statistics, time series analysis, optimization, and simulation to deploy appropriate models for analysis and prediction.

DSDA03 Identify, extract, and pull together available and pertinent heterogeneous data, including modern data sources such as social media data, open data, governmental data.

DSDA04 Understand and use different performance and accuracy metrics for model validation in analytics projects, hypothesis testing, and information retrieval.

DSDA05 Develop required data analytics for organizational tasks, integrate data analytics and processing applications into organization workflow and business processes to enable agile decision making.

DSDA06 Visualise results of data analysis, design dashboard and use storytelling methods.

APPENDIX 5

Edison project (Demchenko et al., 2017): Data science data management

DSDM Develop and implement **data management strategy** for data collection, storage, preservation, and availability for further processing.

DSDM01 Develop and implement data management **strategy** for data collection, storage, preservation, and availability for further processing.

DSDM02 Develop and implement relevant data models, define metadata using common **standards** and practices, for different data sources in variety of scientific and industry domains.

DSDM03 **Integrate** heterogeneous data from multiple source and provide them for further analysis and use.

DSDM04 Maintain historical information on data handling, including reference to published data and corresponding data sources (**data provenance**).

DSDM05 Ensure data quality, accessibility, interoperability, compliance to standards, and publication (**data curation**).

DSDM06 Develop and manage/supervise policies on **data protection**, **privacy**, **IPR** and **ethical issues in data management**.

APPENDIX 6

DAMA DMBok2 (DAMA International, 2012)

Data governance

planning, oversight, and control over management of data and the use of data and data-related resources. While we understand that governance covers 'processes', not 'things', the common term is Data Governance, and so we will use this term.

Data architecture

the overall structure of data and data-related resources as an integral part of the enterprise architecture.

Data modeling & design

analysis, design, building, testing, and maintenance.

Data storage & operations

structured physical data assets storage deployment and management.

Data security

ensuring privacy, confidentiality and appropriate access.

Data Integration & Interoperability – acquisition, extraction, transformation, movement, delivery, replication, federation, virtualization and operational support.

Documents & content

storing, protecting, indexing, and enabling access to data found in unstructured sources (electronic files and physical records), and making this data available for integration and interoperability with structured (database) data.

Reference & master data

managing shared data to reduce redundancy and ensure better data quality through standardized definition and use of data values.

Data warehousing & business intelligence

managing analytical data processing and enabling access to decision support data for reporting and analysis.

Metadata

collecting, categorizing, maintaining, integrating, controlling, managing, and delivering metadata.

Data quality

defining, monitoring, maintaining data integrity, and improving data quality.

Resources

In addition to the extensive references quoted in the report, of particular value for information professionals interested in this area, the following introductory learning resources are recommended:

Stuart, D. (2020). *Practical Data Science for Information Professionals*. Facet Publishing.

Foster project, www.fosteropenscience.eu.

The carpentries, carpentries.org.

AI4K12, github.com/touretzkyds/ai4k12/wiki.

Alan Turning Institute, www.turing.ac.uk.

ADA Lovelace Institute, www.adalovelaceinstitute.org.

Case study sources

The Health Service Executive's LAMA (p 22)

Laura Rooney Ferris, Library Resources Manager; Digital Knowledge Service, Health Service Executive (HSE), National Health Library & Knowledge Service, laura.ferris@hse.ie.

Roche Products Limited: LibraryAnn (p 22)

Arthur Robbins, Library, Information & Knowledge Services Manager, Roche Products Ltd, arthur.robbins@roche.com.

Create Crates (p 25)

Jonathan Clayton, Library Officer, Wakefield Libraries, jclayton@wakefield.gov.uk.

CILIP: the Chartered Institute of Library and Information Professionals

7 Ridgmount Street
London WC1E 7AE

cilip.org.uk