This is a repository copy of *The ethics of AI for information professionals: eight scenarios*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/187487/

Version: Accepted Version

The Ethics of AI for information professionals: eight scenarios

Andrew Cox

The Information School, University of Sheffield, UK

a.m.cox@sheffield.ac.uk

Andrew is a Senior Lecturer in the Information School, Sheffield. Much of his research relates to the information professions and their response to contemporary trends such as artificial intelligence, datafication and managerialism.

The Ethics of AI for information professionals: eight scenarios

Artificial Intelligence (AI) is central to transformative changes happening in many industries, perhaps potentially to a fourth industrial revolution, but it has also raised a storm of ethical concerns. Information professionals need to navigate these ethical issues effectively because they are likely to use AI in delivering services as well as contributing to the process of adoption of AI more widely in their organisations. Professional ethical codes are too high level to offer precise or complete guidance. In this context the purpose of this paper is to review the relevant literature and describe eight ethics scenarios of AI which have been developed specifically for information professionals to understand the issues in a concrete form. The paper considers how AI might be defined and presents some of the applications relevant to the information profession. It thensummarises the key ethical issues raised by AI in general both those inherent to the technology and those arising from the nature of the AI industry. It considers existing studies that have discussed aspects of the ethical issues specifically for information professionals. It then describes a set of eight ethics scenarios that have been developed and shared in an open form to promote their reuse.

Keywords: the information profession; ethics; artificial intelligence; data; social justice

Introduction

A code of ethics is part of what differentiates professions from other occupations. The claim that professions make for autonomy in decision making based on their expertise implies that professionals also need principles of good conduct to protect clients and the public. Professional ethics codes enshrine these principles to guide decision-making in dilemma situations and to underwrite accountability. They are also used in shaping the identity of the profession as a community (Frankel, 1989).

New technologies often seem to pose novel challenges to professional ethics and values, though whether they change issues fundamentally is more debatable (Ferguson,

et al., 2016). One such technology that seems to raise a host of new ethics issues is artificial intelligence (AI). In the last five years there has been much excitement around AI in many realms of activity. But some notable failures have also triggered considerable concern around the ethics of data and AI. In response there has been a proliferation of ethics principles published, though critics still find that these are not enough to ensure ethical practice, suggesting the need for more reflection on the underlying issues. Most domains of work and life seem to be being impacted by AI, including many applications relevant to information professionals. So, it is important for information professionals to think how their existing ethics and values apply to the new technology.

One means of doing this is through the discussion of ethics scenarios (Bradshaw, 2012). Scenarios make abstract principles more tangible and meaningful (Ferguson, et al., 2016). They are often used in training and teaching. In this context, the purpose of this paper is to review the literature of the ethics of AI and to describe eight ethics scenarios of AI which have been developed specifically for information professionals. The questions guiding the paper are:

- What ethical issues does AI pose as a technology and as an industry?

- What specific issues in information work does it pose?

- How can these be captured as ethics scenarios?

- The paper is laid out as follows: First the value of scenarios for professional ethics is established through a brief reflection on the importance and character of professional codes of practice. Then the paper offers an analysis of what AI is and particularly how its application might impact the work of information professionals. The paper summarises both the key ethical issues of AI as a technology and then issues arising from how the industry is organised. It then

reviews literature specifically from the information professional domain. Building on this, the paper goes on to describe the scenarios that have been developed by the author.

The role of professional ethics

A profession is a "moral community" (Frankel, 1989:110). It shares values and ethical principles. A profession's ethical principles offer public and client accountability in the context of professional autonomy, they also help socialise new entrants into the profession and build a professional sense of community. But because they serve multiple audiences they can be hard to create. Indeed, it has been suggested that the process by which an ethics code is developed is as important as the code itself (Frankel, 1989). Frankel (1989) suggests that ethics codes can be of three types (though actual codes may contain elements of all three):

- aspirational (statement of ideals)
- educational (contains much supporting detail explaining the code) and
- regulatory (designed to govern conduct).

Like other professions, information professions have codes of ethics, though there is not international consensus on what they should contain or how presented.

Table 1 identifies the ethics principles mentioned by more than half of the 28 national LIS ethics codes as identified by Shachaf (2005). It reveals a stress on professionalism (such as professional development and integrity) but does identify key ethical principles oriented to the public: confidentiality, access to information and freedom from censorship. There is considerable diversity across the professional community, but a core of agreement (Foster and McMenemy, 2012).

Table 1 Frequently mentioned topics in national LIS ethics codes (data derived from Shachaf, 2005)

Issue	Frequency of mention
Confidentiality and privacy	24/28 (86%)
Professional development	23/28 (82%)
Integrity	19/28 (68%)
Conflict of interest and personal gain	19/28 (68%)
Free and equal access to information	17/28 (61%)
Censorship	16/28 (57%)
Responsibilities to profession	15/28 (53%)

Ethical codes in the information domain tend to be short (Shachaf, 2005). They increasingly seem to stress positive values as much as ethical obligations, suggesting an emphasis on aspiration rather than education or regulation. CILIP's (2018) code of ethics a good example in this context.

The short length of codes such as the CILIP principles make them memorable and powerful, but also reinforces the complimentary value of scenarios to help practitioners work through issues in specific contexts. Such codes are inspirational but they are not much immediate help in navigating complex ethical issues such as raised by AI. Carter (2020) has sought to map CILIP's ethics code to the process of machine learning, and the impression is that many of the principles apply at all stages. However, a more fine-grained approach is needed to articulate the specific challenges around AI in enough detail to guide action in dilemma situations. Some information professional work creates dilemmas when different principles are in conflict (Ferguson, et al., 2016), so scenarios that surface these tensions are useful. They stimulate discussion through which solutions can be found collectively.

Furthermore, Ferguson, et al.'s (2015) study of ethical awareness around RFID suggests that where technologies are concerned perhaps information professionals are over-reliant on vendors to ensure the technology is ethical. Given the vested interests suppliers always have in promoting their products, it is essential that practitioners apply their ethics principles independently. This reinforces the need for scenarios to prompt information professionals to think through the ethics issues around AI drawing on first principles.

Definitions of AI and its role in the work of information professionals

Defining AI is itself complex and contentious, complicating the task of identifying the ethical issues in building scenarios. Firstly, there seem to be multiple specific technologies that make up what is currently being labelled AI. Thus Gartner (Lowendahl & Calhoun Williams, 2018) identity "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. McKinsey's (2018) suggest that "AI capability" consists of: 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. Quite apart from the difference between the two lists, each of these technologies might pose specific ethical dilemmas.

Secondly, it is also important to acknowledge a strong link to issues around data (including big data). Many of these technologies, particularly those listed by Gartner, are about enabling more forms of input as data e.g., speech recognition or computer vision. Equally we know that machine learning relies on large amounts of training data

(and computing power to process it). So previous debates about the ethics around data, such as privacy, consent and transparency are also fundamentally relevant to AI (Zook et al., 2017; Zwitter, 2014).

Thirdly, there is the impact on jobs. Much of the debate around AI revolves around how it will impact labour. AI has the potential for a range of impacts on jobs in replacing, complementing, dominating, augmenting, dividing or rehumanizing work (Global Partnership on AI, 2020). Perhaps all of these will happen to some degree, in different areas of work, at different times, while some work will remain fundamentally unchanged. But considering the impact on work is an important part of the ethical debate.

Fourthly, technically focussed definitions or even a focus on employment take us only so far. It is also important to acknowledge that AI is a set of commercial applications and a related industrial complex. It is hard to separate AI technologies and the power of giants such as Google and Amazon who are the leaders in exploiting it. Some of the current instantiations of AI are entangled with the power of the big Tech companies. Furthermore, AI is an industrial complex. Crawford and Joler (2018) brilliantly deconstruct the Amazon Alexa dot voice assistant showing how it is only made possible through exploiting users' data, human labour (often exploiting click workers), the extraction of raw materials and by expending a vast amount of energy. Machine learning is energy hungry (Brevini, 2020; Lucivero, 2020). Environmental impact audit becomes increasingly relevant to AI. Again, this may not seem immediately relevant from the practical point of view of applying AI to information work but should expand our view of the ethical implications. It is not just about what the technology can do, we should consider the power structures within which it is being created and its human, societal and environmental implications.

On the one hand, this complexity could invite us to simplify the issues to relate to one technology. This is the approach usefully developed by Cooke where she relates a typical machine learning process to the skills and ethics principles of information professionals (2020). However, this limits our engagement with the full complexity of the issue. Thus, in the scenarios developed here an attempt is made to reflect some of the range of technologies, data related issues and the more societal aspects of AI, including its sustainability and impact on jobs.

Equally it is important to stress that the ethics of AI should not be discussed outside the context of its potential benefits. Ethics scenarios must allow the user to consider these. Many applications of AI appear to offer the chance to save lives, improve well-being or impact positively on sustainability. They certainly promise to increase access to knowledge. Dilemmas often arise from trade-offs between different forms of gain and loss.

While this paper has argued for a broad conceptualisation of AI as needing to underlie our scenarios, they must represent the main applications of AI relevant to information professionals. Cox (2021) offers one summary of some of the main areas of application. More analytically, it might be useful to see AI as relating to the work of information professionals in four ways. Firstly, AI tools might be used directly by information professionals in their work e.g., summarisation or translation tools could support information activities such as current awareness. Secondly, AI could be applied directly to information services. There are many applications in knowledge discovery, for example, such as through new search methods and recommendation. In keeping with our broad definition of AI we would have to recognise that many of the features of google search or amazon recommendation are based on AI. Applying such techniques to library supplied services would improve access within special collections or collections

of published literature. Another example of AI applied to information work would be the use of chatbots to answer user queries. A third and somewhat different type of use of AI could be in decision making in the management of information services e.g., for monitoring, predicting or even nudging user behaviour. In addition, a fourth way that AI might relate to information professionals is that there is an important potential role for information professionals in supporting their organisations to use AI for their other purposes, such as by locating or licensing data and finding or procuring systems, understanding IPR issues, or through promoting data and AI literacy. The intention of the scenarios is that they cover the range of these roles.

The ethics of AI technologies

There is no doubt that the possibilities of AI in information work and in other domains are exciting, but they are certainly controversial. Jobin et al. (2019) identified no fewer than 84 ethics codes published by international bodies, national governments, tech companies and others seeking to provide guidance on what ethical AI would look like. A number of authors have attempted to synthesise the core ethical issues from the proliferation of statements. For example, Fjeld et al. (2020) identify eight themes found in 36 ethics guidelines: Privacy, accountability, safety and security, transparency and explainability, fairness/ non-discrimination, human control of technology, professional responsibility, and promotion of human values. Jobin et al. (2019) suggested there was convergence around five main themes in the 84 codes they examined: Transparency, justice and fairness, non-maleficence, responsibility, and privacy. However, a number of other issues such as sustainability and solidarity are also present, if less frequently (Jobin et al., 2019; Hagendorff, 2020). Floridi and Cowls (2019) draw out five core principles: beneficience, non-maleficience, autonomy, justice, and explicability. The

latter refers to AI being both intelligible and that there be accountability for its decisions.

For the purposes of analysing the development of the scenarios, it is useful to offer a short summary of the main issues from the author's own analysis:

- Bias. A critical issue is the way that AI can lead to biased outcomes. This might be because the algorithm itself instantiates bias, or because the data to which it is applied is biased. If historic data reflects past inequalities their use by AI will be likely only to reproduce inequality.

- Transparency, explainability and accountability. It is hard to explain decisions arising from machine learning since these are based on the algorithm itself discovering patterns in data, rather them being written by a human into the code. This makes AI systems opaque. It leads to a challenge of explainability (impacting the ability to give truly informed consent) and of transparency, with the risk of deresponsibilisation: Who is responsible for an error made by the computer? An AI system might sit almost invisibly within a network of systems, so how can a citizen even know that a decision is being made using AI and that it might be based on data that is misleading or inaccurate?

- Privacy. Storing and exploiting data from users is central to AI but creates a risk to privacy. Aggregating different forms of data, even if they are anonymised, can make it possible to infer identities.

- Safety and security. As dependence for decision making on data increases so the risk of systems being compromised becomes critical.

- Human choice. The decision-making potential of AI inherently challenges human agency, particularly when further linked to the issues of explainability.

The ethics of the AI industry

These issues seem to be inherent to the technology so need to be built into any ethics scenarios. But further, perhaps less obvious issues arise because of the way the AI industry is organised. While many governments and other organisations have published codes of principles for ethical AI, critics have suggested that they are not very successful in achieving ethical practice for a number of reasons. This relates strongly to the point already made about the nature of AI not simply as a bundle of technologies, but as a set of commercial applications and a related industrial complex.

A common criticism is that existing ethical codes are too abstract to be useful in practice (Green, 2021a; Haussermann and Lutge, 2021; Mittelstadt, 2019). This reinforces the value of scenarios which can help people think through the issues in a specific context. AI ethics codes tend to focus purely on the design process (not the wider context of decision making) and place the burden of decision making on individual technical experts. As Abbott (1983) points out professional ethics codes focus on individuals not organisations, so evade the question of organisational responsibility (Haussermann and Lutge, 2021). This suggests that the scenarios being developed here must encompass wider contexts for the decision making around AI.

Critically, commentators have argued that the development of AI is dominated by corporate logics. Often ethics statements may merely be "ethics washing" in response to public concern and motivated by the desire to avoid regulation (Metcalf et al., 2019). Indeed, while design decisions are seen as requiring ethical review, business practices driving the decision making are often not (Greene et al., 2019). In reality, ethically developed code could still be applied unethically (Metcalf et al., 2019). The ability of corporations to do this may reflect the relative weak development of data science as a profession. It seems clear that data science is in the early stages of developing as a profession and until it becomes more mature, having its own code of

ethics and collective understanding of how it is to be applied, it is unlikely to provide the guidance designers need particularly in the face of corporate power (Mittelstadt, 2019). Furthermore, the nascent profession is itself not representative of the whole of society. One reason AI might be biased, is because of structural inequalities in the industry itself. It is not surprising that AI is biased when women, people of colour, people with disabilities and older people are under-represented among those designing it. Both points suggest the need for a wider range of stakeholders to be involved in AI ethics as a participatory process, a dimension of the issue that should appear in the scenarios (Greene et al., 2019; Haussermann and Lutge, 2021).

Many critics insist on the need to differentiate an application of AI that is socially neutral and one that promotes valued ends, such as social justice (Green, 2021b). It is not enough in an unequal society for technologies to be fair within the status quo, they argue. Furthermore, for more positive change to be enabled to happen through technologies there must be an understanding of how injustice such as inequality occurs and is sustained. Most computing education does not engage with such essentially sociological issues, so attempts to apply computers to improve society lack a sophisticated analysis of social problems. Rather it is argued, thinking about technology tends to be developed with technological deterministic and solutionist assumptions, that portray technology as a magic fix to complex social problems and as inevitable vehicles of progress (Green, 2021a; Greene et al., 2016). These certainly seem to be part of the Silicon Valley ideology around technology, but they are problematic views. Indeed, many authors see the enthusiasm for AI as linked to trends such as datafication, metrification, dataveillance and commodification, in other words to an increasing stress on measuring human behaviour, increased surveillance of human behaviour, and the wider commercialisation of life. Some critics have suggested similarly that LIS tends to

treat technologies through a technocratic ideology (Mirza and Seale, 2017). In order, for our scenarios to engage with these issues there must be prompts to reflect on how social justice is impacted by AI, informed by a sociological understanding of the structures that maintain inequalities. It is essential that applications of AI fit the culture of the organisation and fix problems in appropriate ways driven by an analysis of the problem, not driven by an assumption that technology is always the right solution (JISC, 2021).

LIS specific ethics issues already debated

Many of the ethical issues with AI discussed in the previous two sections are relevant to information professionals, but there are areas where there might be particular concerns, notably in the case of when AI is applied directly in information services. Before considering these issues, it is worth pausing again to acknowledge the many potential benefits AI may bring in the context of information professional work, such as to make processes more efficient but also to make content accessible on scale (Cordell, 2020), to offer more personalisation and adaptivity of information, and potentially vastly increase access to information. Given information professionals' universal focus on access to information these benefits are ethically significant.

IFLA (2019) provides a good presentation of some of the main ethical issues arising when AI is being applied in the information environment. Many relate specifically to aspects of access to information and freedom of expression, aspects not as prominent in the wider debate on AI ethics, suggesting a distinctive focus for information professional concerns. Thus IFLA (2019) explain that there is the risk that, despite its potential benefits, personalisation leads to echo chambers or filter bubbles, where recommendations reinforce existing beliefs, nudge behaviour or even seek to create addiction (Milano et al., 2020). This effectively limits free access to information. AI's use in forum moderation could also limit freedom of speech, particularly where it

is set up to err on the side of caution (Privacy International, 2018). AI is also a threat to trust in information through its use in the creation of deepfakes. An important concern from an information professional perspective is that lack of privacy can create a "chilling" climate in which people do not feel safe to search for and read the information they want to. This echoes previous debates about the risks of surveillance and dataveillance (Privacy International, 2018).

The points made by IFLA (2019) reflect continuity in debates about AI from previous ethical dilemmas in the profession. Another pre-existing debate that could help us grasp the dimensions of ethical issues of AI in information work is that around learning analytics, which are data about learning behaviours including library use (Jones and Salo, 2017; Jones et al., 2020). Since AI is based on data many of the same challenges exist with AI, at least in an educational setting. Indeed, many of the issues identified echo those summarised above around privacy; safety and security; and transparency, explainability and accountability. But the work around the ethics of learning analytics for information professionals opens up further dimensions of the issues.

Gathering and connecting data about individuals obviously creates a risk to privacy and a danger of reidentification when anonymised datasets are connected. Furthermore, security risks are heightened when data about individuals is amassed. Data accuracy and integrity becomes hugely important if decisions are being made on the basis of it. It is also suggested that there is a threat to intellectual freedom arising from monitoring of user learning behaviours and the use of techniques to nudge behaviour.

Learning analytics projects reveal some troubling patterns about how libraries have been tempted to use data about users (Jones et al., 2020). Often learning data was being used without consent; certainly, students were often unaware of how their data

was being used. It also appears that often there had been no ethics review. Libraries

rarely had published policies on responsible use of the data. Further, there was a lack of

clarity about what or who benefits. Although it is students' data it could be the

institution benefiting more than the students. The stress on the benefits being to learners

in educational applications of AI and that it promotes equity by Institute for Ethical AI

in Education (2021) is compelling. Another point made by Jones et al. (2020) is that in

the context of joint projects, there was felt to be a risk of librarians' ethics concerns

being over-ruled, especially as user data comes to be seen as an organisational asset.

These issues are relevant to information professionals' role in AI projects in their

organisation.

Ultimately learning data consists of what can be measured and is not often

actually direct measurement of learning. Taking a book out of the library, as a learning

datum, might correlate with learning, but really tells us nothing about what has been

learned. This issue of validity has important implications for AI which has tended to

exploit the data to hand, without sufficient consideration of the provenance, meaning

and limits of data. The debate around the ethics of learning analytics reveals some of the

risks associated with libraries engagement with AI.

Method of developing the scenarios

Following the discussion presented above, the scope in developing the scenarios was to

encompass:

(1) The potential benefits of AI as well as the ethical challenges;

(2) The wide range of possible applications of AI, and also the view of it as an

industrial complex;

(3) The range of AI applications relevant to information professionals, from direct uses in information services to its use in service management and the case where information professionals might be supporting organisations to deploy AI;

(4) Applications in the different contexts of information work, e.g. health, libraries, commercial and legal sector etc;

(5) Underlying data issues;

(6) The full range of ethical dilemmas that the wider literature throws up, including bias; transparency, explainability and accountability; privacy and consent; safety and security; and human choice, but also wider issues around participation; its impact on social justice and sustainability and the links to issues such as datafication and commodification.

(7) The impact on jobs;

(8) The whole decision-making process around application of AI, not merely in the development process;

(9) The ethical concerns around access to information and freedom of expression of particular concern to information professionals, such as the "chilling" effect of surveillance.

In preparing the first draft of the scenarios reference was made to guidance set out in the Institute of Business Ethics good practice guide (Bradshaw, 2012) such as: 1) ensuring, as far as possible, that the scenarios place the dilemmas in relatable contexts that reflect the cultural contexts of information work. 2) Avoiding stereotypes in the scenarios that themselves can pose ethical issues. 3) Matching to the audience. 4) Being stylistically open, rather than implying a simple, reductive moral. 5) Using a second person narrative point of view which increases identification and immediacy by placing the reader in the position of the person with the dilemma. 6) Keeping the scenarios as

concise as possible and avoiding excessive detail. The purpose of the scenarios is to stimulate debate: So, it is not intended that there is a correct answer, rather a section of guidance notes and questions was created to accompany each scenario. To ensure their on-going value it was decided to share them on a CC-BY-SA licence, allowing them to be adapted for local use or updated as perceptions of AI change.

After developing a collection of eight scenarios (DOI 10.15131/shef.data.15147411.v1), a number of experts were recruited from across the profession and asked to offer informal comments on the first draft and on AI ethics in general. Their input was used to produce a second version of the scenarios (DOI 10.15131/shef.data.17081138.v1). The major changes were to elaborate the notes extensively to draw out the issues further.

The scenarios

The scenarios produced through this process can be accessed at DOI https:/doi.org/10.15131/shef.data.17081138.v1 . A title and brief description of the scenarios would be as follows:

(1) Supporting first responders – data managers raise objections to sharing data for a system to support

(2) Nudges – library asked to contribute data to a tool that nudges students to improve their well-being

(3) The voice assistant – A library offers a voice assistant service to the public

(4) A special collection – A donation is predicated in enabling access to controversial content

(5) Forum moderation – automation of moderation of public forum

(6) The recommender system – responses to a new recommendation tool that makes recommendations of material based on past reading

(7) Stakeholders – involvement of stakeholder communities in an AI project

(8) Project partners – concerns about power and ethics in a joint project

Table 2 offers a high-level summary of the scenarios, identifying their topic, field of application, the type of AI, how information professionals were imagined to be involved and the ethics issues raised.

Table 2. Summary of the scenarios

Title	Field of application	Type of AI	Information professional involvement	Issues raised
1. *Supporting first responders*	Health	Advice system	Data supply and quality	Consent Privacy Security Transparency Commodification
2. *Nudges*	Education	Nudging tool	Library supplying data	Human agency Privacy (Fine grained) consent

3. *The voice assistant*	Public information	Voice assistant	AI provides a question answering service	Educating users/ giving information Jobs Bias Sexism
4. *A special collection*	Archive	Search tool	AI enables access to content	Bias Access to information
5. *Forum moderation*	Public information	Forum moderation	AI moderates forum content	Bias Access to information Freedom of expression
6. *The recommender system*	Academic	Recommendation	AI provides recommendations to users	Access to information Chilling effects Lack of transparency Information skills Privacy Bias

				Commodification
7. *Stakeholders*	Any	Any	AI project lead	Data bias Representation Image/ expertise of profession
8. *Project partners*	Any	Any	AI project participant	Environmental impact Commodification Representation Image of organisation/ profession

Scenario 1, entitled Supporting first responders, places the reader in the situation of being an information professional coordinating access to data for an AI tool that helps first responders improve treatments in an emergency. It is intended to throw into relief the tension between highly beneficial, lifesaving applications which nevertheless raise a range of issues, such as privacy, consent, transparency, and security. As with each scenario, the reader is invited to weigh up what the ethics issues are and what action they might take. As the notes explain, legal issues could interact with the ethics issues. Scenario 2, Nudges, poses the dilemma of the library being asked to supply to data to support a well-being App. Again, there are acute issues of privacy and consent.

Can nudging behaviour be justified, even if there are health benefits to the person whose behaviour is being influenced? How is the issue of consent to be handled? Such questions surface the issues debated around learning analytics, reviewed above. The most fundamental question is how this technical solution fits into wider strategies to support student mental health and well-being. This should not be a technological solution in search of a problem, the outcome of a holistic analysis of the problem that has identified AI technology as the best available approach.

In scenario 3, the voice assistant, a public information service is using AI to answer public queries. There are issues around accessibility and equality of service, but also a potential impact on jobs and also human relations in the reference encounter. Scenario 4, a special collection, imagines the situation where a donation to support acquisition of AI tools is linked to what could be deemed controversial content. Scenario 5, forum moderation, imagines the case where a public information forum needs automated moderation, but this risks blocking content that is legitimate. There is also the issue of transparently explaining how the system works, since like much AI it is based on machine learning. Scenario 6 is called the recommender system. Various responses to the recommendations are reported prompting questions about how the system needs to be configured and explained to ensure consent and transparency and to avoid chilling effects of apparent surveillance.

Scenarios 7 and 8 are rather different. Rather than focussing on the dilemmas around a specific AI application they present dilemmas around the development and implementation process that apply generically regardless of the specific application. In scenario 7 the challenge is of fully involving community stakeholders is balanced with professional expertise in user need. Scenario 8, project partners, imagines an information professional joining a larger project and developing concerns about how

ethics are being handled across the project. Like scenario 1, this reflects the likelihood that information professionals will be involved in wider AI projects within organisations not just applying it directly to information work.

Conclusion

AI is likely to play a major part in the digital transformation touching information work as it will in many other sectors. AI promises to greatly increase access to knowledge, but it also carries with it complex ethical challenges. Information professionals need to have an understanding of the issues to enable them to develop responsible operations in their own services and to support ethical use of AI in their wider organisation (Padilla, 2019). This paper contributes to informing and educating information professionals by describing the development of a set of scenarios instantiating the key ethical challenges set in relatable information service contexts. The paper analyses the main ethical issues raised by AI and draws on the wider literature that explores the barriers to ethical practice of AI. It also reflects on information professional specific issues, such as the chilling effects of AI related surveillance, and on previous debates, such as around learning analytics, that shed light on the particular challenges in the information professional context. A description of the final scenarios is presented, with free access to the scenarios being available on a CC-BY-SA licence at DOI https:/doi.org/10.15131/shef.data.17081138.v1 . By making the scenarios available on this licence it is hoped to encourage their wider use. They can be adapted for local conditions and updated as our understanding of AI evolves.

References

Abbott, A. (1983). Professional ethics. *American journal of sociology*, 88(5), 855-885.

Association of Research Libraries (2019). *Ethics of Artificial Intelligence.* Association of Research Libraries. https://doi.org/10.29242/rli.299

Bradshaw, K. (2012). *Developing and using business ethics scenarios*. Institute of Business Ethics.

Brevini, B. (2020). Black boxes, not green: Mythologizing artificial intelligence and omitting the environment. *Big Data & Society*, 7(2), 2053951720935141.

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.

CILIP (2018). Ethical framework. CILIP. https://cdn.ymaws.com/www.cilip.org.uk/resource/resmgr/cilip/policy/new_ethical_framework/cilip_s_ethical_framework.pdf

Cordell, R. (2020). *Machine learning and libraries: a report on the state of the field.* Library of Congress. https://apo.org.au/node/307049

Cox, A. (2021). *The impact of AI, machine learning, automation and robotics on the information professions: A report for CILIP* https://www.cilip.org.uk/general/custom.asp?page=researchreport

Crawford, K. and Joler, V. (2018) *Anatomy of an AI system*, https://anatomyof.ai/

Ferguson, S., Thornley, C., & Gibb, F. (2015). How do libraries manage the ethical and privacy issues of RFID implementation? A qualitative investigation into the decision-making processes of ten libraries. *Journal of librarianship and information science*, 47(2), 117-130.

Ferguson, S., Thornley, C., & Gibb, F. (2016). Beyond codes of ethics: how library and information professionals navigate ethical dilemmas in a complex and dynamic

information environment. *International Journal of Information Management,* 36(4), 543-556.

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

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*. Berkman Klein Center Research Publication, (2020-1).

Foster, C., & McMenemy, D. (2012). Do librarians have a shared set of values? A comparative study of 36 codes of ethics based on Gorman's Enduring Values. Journal of librarianship and information science, 44(4), 249-262.

Frankel, M. S. (1989). Professional codes: Why, how, and with what impact? *Journal* of business ethics, 8(2), 109-115.Global Partnership on AI (2020). Future of work working group report. https://gpai.ai/projects/future-of-work/gpai-future-of-work-wg-report-november-2020.pdf

Green, B. (2021a). *The Contestation of Tech Ethics: A Sociotechnical Approach to Ethics and Technology in Action*. arXiv preprint arXiv:2106.01784.

Green, B. (2021b). Data science as political action: Grounding data science in a politics of justice. *Journal of Social Computing*, 2 (3) 249-265. https://doi.org/10.23919/JSC.2021.0029Greene, D., Hoffmann, A. L., & Stark, L. (2019, January). Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. In *Proceedings of the 52nd Hawaii international conference on system sciences*.

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), 99-120.

Häußermann, J. J., & Lütge, C. (2021). Community-in-the-loop: towards pluralistic value creation in AI, or—why AI needs business ethics. *AI and Ethics*, 1-22.

IFLA (2019). *IFLA statement on libraries and artificial intelligence*. IFLA. https://www.ifla.org/wp-content/uploads/2019/05/assets/faife/ifla_statement_on_libraries_and_artificial_intelligence.pdf

Institute for Ethical AI in Education (2021) The ethical framework for AI in Education. https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf

JISC (2021) A pathway towards responsible ethical AI. https://www.jisc.ac.uk/reports/a-pathway-towards-responsible-ethical-ai

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.

Jones, K. M., & Salo, D. (2018). Learning analytics and the academic library: Professional ethics commitments at a crossroads. *College & Research Libraries*, 79 (3) 304-323.

Jones, K. M., 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. Jones, KML, Briney, KA, Goben, A., Salo, D., Asher, A., & Perry, MR, a Comprehensive Primer to Library Learning Analytics Practices, Initiatives, and Privacy Issues. *College & Research Libraries*, 81(3), 570-591.

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

Lucivero, F. (2020). Big data, big waste? A reflection on the environmental sustainability of big data initiatives. *Science and engineering ethics*, 26(2), 1009-1030.

McKinsey Global Institute. (2018). Skill Shift: Automation and the Future of the Workforce (Discussion Paper, May 2018). In McKinsey &Company (Issue May). https://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

Metcalf, J., & Moss, E. (2019). Owning ethics: Corporate logics, silicon valley, and the institutionalization of ethics. *Social Research: An International Quarterly*, 86(2), 449-476.

Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender systems and their ethical challenges. *AI & SOCIETY*, 35(4), 957-967.

Mirza, R. and Seale, M. (2017). "Who Killed the World? White Masculinity and the Technocratic Library of the Future." In *Topographies of Whiteness: Mapping Whiteness in Library and Information Science,* edited by Gina Schlesselman-Tarango, 171–97. Sacramento, CA: Library Juice Press.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501-507.

Padilla, T. (2019). *Responsible Operations: Data Science, Machine Learning, and AI in Libraries.* OCLC Research Position Paper. OCLC Online Computer Library Center, Inc. https://www.oclc.org/research/publications/2019/oclcresearch-responsible-operations-data-science-machine-learning-ai.html

Privacy International (2018). *Privacy and freedom of expression in the age of artificial intelligence*. https://privacyinternational.org/report/1752/privacy-and-freedom-expression-age-artificial-intelligence

Shachaf, P. (2005). A global perspective on library association codes of ethics. *Library & information science research*, 27(4), 513-533.

Thornley, C., Ferguson, S., Weckert, J., & Gibb, F. (2011). Do RFIDs (radio frequency identifier devices) provide new ethical dilemmas for librarians and information professionals?. *International Journal of Information Management*, 31(6), 546-555.

Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S. P., ... & Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLoS computational biology*, 13(3), e1005399.

Zwitter, A. (2014). Big data ethics. *Big Data & Society*, 1(2), 2053951714559253.