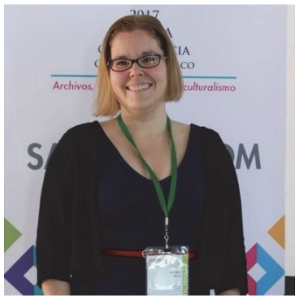
A Brave New World: Artificial Intelligence and Archives



Anthea Seles
Secretary General of International Council on Archives

Summary

Artificial intelligence and machine learning can provide many opportunities for archives and the archival profession. The mass production of data has introduced challenges when it comes to identifying records of long-term value and artificial intelligence offers us a way to automate this process, but it equally bring with it many challenges. During the course of this presentation, we will explore the impact of artificial intelligence use in government, the impact of AI on archival appraisal and selection along with its impact on the access and re-use of archival holdings. The purpose of the presentation will be have the community reflect on the implications of AI as evidence and the ethical questions we need to address if we are to deploy these technologies in our work.

Biography

Dr Anthea Seles, hold a doctorate from University College London (2016), a Masters degree in Archival Studies from the University of British Columbia (2003) and an honours undergraduate degree in Art History from Queen’s University (2001).

As Secretary General, is the senior staff member in the Secretariat, responsible for the line management of all ICA’s paid staff. She co-ordinates the work of ICA officers and bodies on the organization’s major initiatives and manages its relations with external partners.

Prior to joining the ICA Dr Seles worked as the Interim Director of Digital Selection and Transfer at the National Archives UK, where is was responsible for overseeing digital records transfers coming from government departments to the National Archives. During her time there she worked with colleagues to test out machine learning applications to help automate the appraisal, selection and sensitivity review process, she was also responsible for implementing TNA’s first digital records transfer process. She also worked with the International Records Management Trust (2010-2013) on topics such as digital data integrity for development and the importance of records management for accountability and transparency initiatives.

Dr Seles has lectured extensively at a number of international conferences on topics ranging from artificial intelligence, digital preservation and records management for accountability and transparency

A Brave New World: Artificial Intelligence and Archives

Anthea Seles
Secretary General of International Council on Archives

NB: These speaking notes do not constitute the complete oral presentation given at the EASTICA conference. There are examples or experiences that may be shared during the course of the presentation that may not be documented in these notes.

[Slide 2- Overview]

In today's presentation I will be discussing the impact of artificial intelligence and machine learning on archival practice. I think a good place in any presentation is to start off with definitions and so I will try and give you some understanding of what is meant by the term 'Artificial Intelligence' along with other terminology that I will be using frequently during the course of this presentation.

Following that, I will touch on the use of artificial intelligence and some of the challenges that are starting to emerge in trying to automate decision-making processes. I think I should explicitly state that artificial intelligence is a new form of record, evidence of government decisions and transactions that we will need to identify and acquire. In the section dedicated to government use of AI I will be asking you to consider your role in advising departments on the use of these new technologies not only to facilitate decision-making but to ensure that they can be accountable and transparent for those decisions. I think we have an incredibly important role to play not only in identifying records of the future but holding governments to account for how they deploy these technologies.

However, the ability to effectively deploy these technologies so that they can be used responsibly is contingent upon access to accurate information. Government digital record-keeping processes have not been ideal, which means that the information we may want to put into these tools may not necessarily give us reliable results, moreover there is a tendency towards combining information (ex: tweets, datasets etc) from multiple sources in order to train a machine to make a decision, but little consideration is given to whether the information should be combined and then whether it's reliable. Again this creates challenges when trying to determine the accuracy and reliability of the results coming from AI

In this Brave New Digital world we will need to consider the impact of AI on our own practices and I will be focusing on appraisal and selection, and then access and re-use. The volumes of information we are dealing with in digital are vast and our ability to use traditional appraisal and selection techniques is limited, nor can we look at every single file or document. Automation is inevitable and we need to understand how to use these machines, their strengths and weaknesses, and when the human needs to intervene.

The reason we acquire records is not to keep them locked up but to make them available for consultation and re-use for researchers, but these technologies breakdown silos which we have constructed via archival description. The question becomes whether it is in the best interest of the archives to break down these walls, and if it is what are the implications?

[Slide 3- Definitions]

I will use certain terminology during this presentation, based on my experiences working with government digital records and to ensure that we are all on the same page. I would like to define what I mean when I use these terms:

DATA:

Structured data: Information, more often numerical information, put in tabular form to enable quantitative analysis.

Unstructured data: Information consisting of word processing documents, power point presentations, videos, sound records, photographs etc.

ENVIRONMENTS:

Structured record-keeping environments: Environments where documents and data are placed in an ordered fashion to allow for retrieval. Ex: Information management system or shared drives with a unified classification scheme.

Non structured record-keeping environments: An environment where documents and information are not organised and can be comprised of a running sequence of document or a shared drive with no unified classification scheme.

[Slide 4- What is Artificial Intelligence (AI)?]

There is no standard definition for Artificial Intelligence and it is more a supra-term for a whole host of machine learning approaches: predictive coding, neural networks or specific algorithms such as Natural Language Processing, Latent Derichlet Allocation, Latent Semantic Indexing.

There are two generally accepted categories: Supervised and Unsurpervised

Supervised: Requires a human to mark up or compile a homogeneous dataset to train an algorithm to recognise patterns or terms in the data. This process requires a lot of up-front work and also requires you to have some level of understanding of the dataset.

Unsupervised: Data is loaded into the system and without any upfront human intervention, analyses the data and provides result.

[Slide 5- Artificial Intelligence, Machine-learning and Neural Networks]

As already mentioned, artificial intelligence is an all-encompassing definition for any activity where a machine/system takes information (structured and unstructured) to predict an outcome.

Subsets of artificial intelligence, again as mentioned, include machine learning and neural networks. Machine-learning is the process of training a system to ‘learn’ how to make a decision using a pre-tagged data whereas neural networks are trained to identify patterns and classify information. By layering neural networks, you can create deep learning networks that can do things like facial recognition or more complex information processing.

[Slide 7- Archival Considerations]

The presentation will be divided into three sections: 'Impact of Artificial Intelligence, machine learning, and data mining in government', 'Use of artificial intelligence on archival processes' and 'making records accessible and readable'

The section on 'Impact of Artificial Intelligence, machine learning and data mining in government' will touch on government use of artificial intelligence, challenges and opportunities and the role of the archivists in this space and why we should care about what is happening.

'Use of artificial intelligence on archival practice' deals directly with the realities we need to face around the digital deluge of data and legacy digital information that still exists in government departments and in private organisations, how poorly this information is organised and managed, along with the concomitant impact this has on the use of artificial intelligence. There are important questions we need to ask ourselves to understand how to deploy these technologies to best effect or else we can incur serious risks.

Finally, in 'Making records accessible and readable' I will consider the effects of new digital research methods being used by digital historians, social scientists, digital humanities and many more, and how these new research practices may challenge what we make accessible and how. Moreover, there are some important questions we need to start asking ourselves about what we digitise, given the advent of machine-learning.

[Slide 8- Government Use of Artificial Intelligence and machine learning 1]

More and more governments are using artificial intelligence to help them make decisions; policy decisions, resource allocations and much more. We have only to look at robotics, immigration processes and security measures to see how machine learning and neural networks can be used, but its use goes far beyond what we can see.

These techniques are being used by data science and statistical units in governments and private corporations. There are very successful ad campaigns developed in the United Kingdom based on data 'scraped' (it means extracted or taken) from social media platforms and put through proprietary algorithms. There government policies today that are developed with the help of artificial intelligence because the mass of data that needs to be analysed is too much for one department, unit or individual to analyse. There are many promising uses for AI but we need to proceed with caution. Too often the technology 'hype' and trying to 'stay ahead of the curve' overtake any ethical considerations when it comes to the deployment of these tools.

[Slide 9- Government Use of Artificial Intelligence and machine learning 2]

The core to any AI project is: What question are you trying to answer? Do you have access to the data you need to try and answer the question? If not where is the data and how do you get it?

However, there are challenges with the data science and other approaches when it comes to the use of machine-learning and AI algorithms in government decision-making: Is the data we are combining meant to be combined? Are we simply comparing apples and oranges?

AI needs data, a lot of data, in order to have enough information to be able to accurately identify patterns to arrive at a decision, but what is not often asked when data is being input into an algorithm is whether the datasets should be combined. If you are studying breast cancer and want to determine the best place to allocate needed medical care you should ensure that the datasets are consistent, in other words you do not want to be combining a dataset of women who have colon cancer and breast cancer, or a dataset that represents a demographic of women between 45-60 from Tokyo, and then another that is purely looking at women between 30-40 from Aomori. The machine's decisions will be extremely skewed and will not reflect the realities on the ground, leading to poor decisions on resource allocations for breast cancer care. This is of course an extreme example, but it is to illustrate that combining datasets without considering the sample demographic and the context of creation will have a noticeable affect on the AI outputs.

Another question that needs to be examined when using AI is whether the data is biased and how does that affect the output of the algorithm? How does that affect what we see and how we interpret it? I will discuss this later in much greater detail using a US study on recidivism rates, but it is worth noting that understanding the context of creation for a training data set for an AI algorithm is essential. We need to ask ourselves hard questions about whether the data is skewed to represent inherent social or socio-economic biases. If this is not taken into account AI used to make things 'fairer'or processes easier can in actuality reinforce social injustices, inequality and poverty. I would recommend you read Cathy O'Neill's Weapons of Math Destruction to understand what occurs when algorithms are implemented without consideration.

Archivists have an important role to play here, we advised organisations on the creation and preservation of records and data to ensure their evidentiary value overtime and when it comes to AI use in government and elsewhere we have an equally vital role to play, but let me ask a few pointed questions:

- What advice would you give on the creation and preservation of 'algorithmic/computational records'?
- Does the archivist have a role to play in advising how algorithms and code are created for decisions-making? How do we know what to preserve and how?

For the first question I think there is much work we still need to do, to be able to advise on the preservation of 'algorithmic/computational records'. It is not insurmountable and there are many archival programmes that are converting themselves into 'computational archives' programmes, teaching upcoming professionals basic coding and other necessary skills to work in a fast-paced digital environment. The British Library has created the Digital Scholarship Training Programme (https://www.bl.uk/projects/digital-scholarship-training-programme) to offer training to existing professionals '…the space and opportunity to delve into and explore all that digital content and new technologies have to offer in the research domain today'. There are also many great online and free courses that archivists can take on 'Introduction to Statistical Methods' and computer coding that can help them build their skills. We can do this we just need to take the time.

On the second questions, we absolutely have a role to play because algorithms are records, not of tomorrow, but of TODAY! I think where we need further reflection is trying to understand what needs to be preserved and how. This is not so simple, because an algorithm is simply a piece of code/rules for a computer to interpret information and so to apprehend how an algorithm arrived at a decision requires ALL the data that was used to train the algorithm, code books, tracking logs etc. There are huge amounts of data and information that need to be captured and preserve, which poses serious infrastructural questions for archival institutions. We would essentially need to set up vast data centre and I don't believe every archives has the resources to do this, as such we need to consider if there are more cost-effective and efficient ways to preserve this large amount of evidence.

[Slide 10- Government Use of Artificial Intelligence and machine learning 3]

Here is an example of a visualisation, it's not a policy visualisation used for government decision making, it is actually a visualisation from JurisData France about divorce proceeding payouts by gender along payments are being appealed. But for a moment let us consider that this is a policy visualisation, in other words it is data combined by government department and graphically represented to help them make decisions about staffing in family courts in different regions.

So one of the first pieces of information that data scientists and policy analysts will want to look at are rulings, which are contained in word documents; unstructured data. This information must be compiled into tabular form along with other data. During this process, decisions will be made about what information will be transferred into tabular form and what will be omitted. Next there may be datasets exported from the case management system that documents number of cases heard in a day, amount awarded, date of the decision, whether the decision was appealed etc. Again, it will be up to the policy analyst or data scientist to document what information from the case management system will be used for the master data for the the policy visualisation. There may also be data from social media used to provide more anecdotal information about experiences in family court.

Throughout this process decisions are being made regarding what information or data points to use or omit, in order to arrive at decision and all this needs to be documented. Including the development of the code that will run over the data to provide results. Code development/Algorithm development is an iterative process and each change should be documented as well, though this is not always the case in reality.

Every phase of the policy visualisation development there are decisions being made, decisions that can affect the output from AI and which can in turn affect people's lives. Our role as archivists is document government and organisational decision and transaction, but also we are stewards of the evidential value of 'records', which by extension – I would argue – means that we have a role to play in ensuring the ethical deployment of these applications in government.

[Slide 11- Government Use of Artificial Intelligence and machine learning 4]

Here is an example government use of artificial intelligence in government from Sky News in the United Kingdom from December 2017 discussing the use handwriting detection to try and identify benefits fraudsters, so people claiming social benefits illicitly.

I think there are several points to consider when looking at the article

1) Benefits claimants are often some of the most vulnerable individuals in our society and whilst receive public funds does demand some accountability on their part, how many of them consented to having their handwriting, which is identifiable personal data used in this way?
2) Automating handwriting analysis is still in the early research phases with promising results coming out of some computer science and digital humanities programmes, but the results are contained to a small controlled sample of data. The biggest challenge with handwriting analysis

is that writing is individual to every person and can change over time. It is not regular/standard which makes it difficult to train a machine and get a set of consistent and reliable results.

3) As such the government is using experimental technologies on some of the most vulnerable individuals in society, to counteract fraud

This a good example of government trying to appear to be innovative and acting without really considering the full impact that these technologies can have on the lives of the most vulnerable. What if it's wrong? How do you argue with the results of a machine? Also how do you argue with government when your entire livelihood has been taken away? Is this worth government investment, when you consider tax evasion by large corporations?

Leaving the very serious ethical questions aside, let us examine for a moment the role of the archivist in this scenario: If this becomes standard practice in government and passes into policy how do we begin to advise on what documentation needs to exist? What does integrity and accountability look like in this context? By extension, what do we preserve?

This algorithm as far as we can discern was (I presume) developed by the government department in question, but what happens when a government or organisation uses an algorithm created by a commercial supplier? How do we understand or know what is happening? How can we trust the answer?

[Slide 12- Government Use of Artificial Intelligence and machine learning 5]

This example looks at the use of third-party algorithms in government decision-making along with the impacts of biased data when training an algorithm. The case study in question takes up a chapter in the book Weapons of Math Destruction and looks at the use of AI to predict recidivism in the US criminal justice system.

The system in question was developed by third-party commercial supplier Northpoint to help judges predict whether a defendant was likely to re-offend. The system they developed was called Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS for short. To begin to understand the issue with this system we need to start with the training dataset and a few facts about incarceration rates for visible minorities in the US

- Sentences given to African-American prisoners in the federal system is 20% longer than those given to white convicts for similar crimes
- African-American represent 13% of the population of the United States, but account for 40% of the prison population

That's a high-level overview of the bias that existed in the training data for the algorithm. This may seem like an obvious statement, but algorithms are only tools that process large amounts of data and base their outputs on the contents of the data. If the data is inaccurate it will amplify the inaccuracies, similarly if the dataset is biased it will only amplify the bias, it cannot fix or compensate for poor datasets. Thus the outputs of the system overly penalised African-American and Latino defendants and because the algorithm was proprietary it was difficult to get any level of accountability around the decision-making process because it was argued that revealing the workings of the algorithm would compromise the commercial advantage that Northpoint had when marketing COMPAS.

This is a case where the algorithm is a ‘black box’ and you cannot understand how the data is being processed. This poses a number of issues when start talking about accountability and transparency of government or judicial decision-making processes.

[Slide 13- Why should this matter to you? 1]

As archivists and information manager we need to recognise that algorithms are ‘records’, they are a new form of evidencing government decision-making which are actually being poorly documented or governments are locking themselves into third-party contracts that affect their ability to be accountable and transparent for the decisions these machines may be assisting with. Ostensibly governments need to be held accountable if they use these technologies to make a decision that has an impact on the lives of their citizens. Moreover, the predilection at this point has been to treat machine outputs as above reproach and there has been a reticence to be transparent about the training process, however this tact cannot continue, not when citizens or people’s lives are being affected. As archivists we are responsible for identifying and preserving that information:

- But what should we preserve? All the components that contributed to training the algorithm? (e.g. documents, data, social media information, algorithm and the results)? Only algorithm and the results? This is a question I asked earlier.

To be able to answer these questions requires us to have the capacity and the skills to advise decision-makers in departments and ministries that are seeking to implement these technologies, but are we invited to the table?

Moreover it requires archives to have the requisite infrastructure to ingest, preserve and make accessible this data.

[Slide 14- Why should this matter to you? 2]

We need to be concerned about this because we will be responsible for preserving algorithms in intermediate and historical archives and we are not currently considered stakeholders when it comes to discussions connected to their development and implementation, which means that many things that are needed to prove the authenticity and reliability of these records may be lost overtime. As with most digital records, we need to be involved at the point of creation, but we also need to envisage the fact that we may need to capture these records shortly after creation because government departments tend to have quite a short term view on the usefulness of this material and once the work has been completed. They will put this material aside and will not be concerned with the long-term maintenance of the algorithms, data sets or any ancillary materials. We need to develop the skills and competencies needed to play our role as trusted adviser now, we cannot wait. As I said earlier, as there are programmes available to reinforce the technical capacity of the profession, but we also need to develop our understanding of the ethical implication of these technologies in order to be effective adivsors. Having said that, while we may be able to acquire all the related materials and advise appropriately we need to accept that we many not fully understand how an algorithm arrived at an output for a decision and we need to accept that we will not achieve perfection; we will need to accept good enough.

[Slide 15- Impact of Information Management Practices]

For several years I worked at the National Archives in the United Kingdom, beginning as Digital Transfer Manager and moving to Interim Head of Digital Selection and Transfer. I worked closely with the unit responsible for government information management advice as I needed to understand firstly how to information was managed in government in order to comprehend what to capture and how during the digital transfer process. This led me and a team of colleagues to being studying digital information practices in government for nearly 3 ½ years. In my studies I found that information management systems are not always easy to use, especially early systems as they were quite rigid and not very user friendly. This meant that users tried to find other, easier ways to file their information.

Often users reverted to methods that had less friction like use shared drives. These were used however in parallel with information management systems, resulting in incomplete folders, idiosyncratic naming conventions and significant duplication. Also shared drives had varying levels of structure, some where well organized, others not at all. As newer, better systems were put in place users started using the information management systems, though not always more efficiently. In some instances, shared drives were closed down, in others they remained available to users who continued to use them in tandem with information management systems. All in all this created large stores of legacy data that government departments needed to parse through in order to find what to keep and what to discard, as well as what was sensitive.

[Slide 16- Impact of Information Management Practices- 2]

The study found that for every terabyte of data in an information management system there was at least 25 TB in shared drives, to note that this figure does not include data held in email servers or datasets. Once we accounted for the totality of the information holdings which includes email servers and datasets it added up to over 1.5 petabytes of data that needed to be appraised and selected. So as an illustrations 1.5 PB = approximately 1.5 billion Word documents, on average. This ratio would, of course, vary according to the types of formats you are dealing this.

The problem with trying to appraise this volume of data is that information management teams did not know what was contained in legacy data holdings and by extension did not know what documents or data needed to be preserved. This information could also have differing levels of contextual information and limited metadata. Metadata could also be compromised because of previous migrations.

We often had to revert to contextual information about the dataset such as timelines to indicate when major events would have occurred and interviews with former or long-standing staff members to create a profile of what the legacy holdings might contain.

[Slide 17- Impact of Information Management Practices on Appraisal and Selection]

The contextual information only helped a little because we needed some way to see inside the dataset, but this was complicated by the large volume of unstructured data that was minimally organised. As such we decided to start a second study to examine off the shelf systems that had basic machine learning capabilities for the purposes of assessing their viability to carry out appraisal and selection, along with sensitivity review. See: The Application of Technology Assisted Review to Born-Digital Records Transfers, Inquiries and Beyond.

http://www.nationalarchives.gov.uk/documents/technology-assisted-review-to-born-digital-records-transfer.pdf

[Slide 18- Artificial Intelligence and machine learning in Records Management and Archives]

What we found is that the systems we looked at were good at basic search such as keyword, Boolean, regular expressions and they could process large volumes of data. These systems had good precision and recall when it came to personal information such as name, social insurance number, phone number and address. Where the systems did not fair well is understanding context and inferring meaning or around handwriting analysis.

It is important to understand when we discuss the use of artificial intelligence and machine learning that humans are not irrelevant in this process. There are things that humans can do better than machines such as understand context, assess and reason through multiple variables in a split second.

We must understand the strengths of the machine and that of human beings in this process it is not an either/or scenario. There is complimentary.

[Slide 21- Artificial Intelligence and machine learning in Records Management and Archives 2]

We encountered several issues and limits during testing:

1. Lack of understanding regarding the content and the context of creation

   As mentioned, the systems we studied did not have the capacity to understand the content and contextualise the information, although they could identify specific data points such as dates, names and keywords. This meant when we needed a more nuanced assessment of certain records, we needed a human.

2. Corruption or alteration of metadata

   The machines cannot compensate for poor metadata or data, thus if there were issues with the dates of creation or modification, we would not be able to identify material for transfer, nor to assess the closure periods of the records. We needed the contextual information along with keyword searches to help us identify record sets, though this method is not perfect, and you can still miss out on important material that should have been identified for transfer.

3. Difficulty understanding the visualisations generated by the machines

   Each of the machines we worked with had similar visualizations but defined records (video, word, containers etc) in different ways and so you needed to be really careful looking at the visualization and understand how different formats or containers are defined by the systems or else you could accidentally discount something as a record for acquisition

4. Understanding the reliability (precision and recall) of the results and the acceptable level of risk

   If you remember earlier when I discussed the issues of using proprietary software for AI and decision-making, well we encountered similar challenges when working with off the shelf machine learning software. We could never be complete certain about the precision and recall of the results and the software providers would not divulge how their algorithm processed data. This is for very valid reasons that is because these are commercial products and the way the algorithm processed information was their asset. However, it still meant we could never be sure about the accuracy of the output, which made it difficult to assess the ‘success’ of the output.

5. Distrust in technology and the results generated by the systems
   The next two bullet points deal with issues we encountered with our liaisons at government departments. In the first instance there were a gamut of views on the trustworthiness of artificial intelligence outputs. Some of our departments completely distrusted the results of the systems because we could not guarantee 100% accuracy in the results and they had zero risk appetite. Whereas other departments accepted the results of the systems without question.

6. Significant time required to ‘train’ the system, departments wanted something much more automated (i.e. unsupervised)
   The types of artificial intelligence/machine learning systems we were testing were supervised AI algorithms and quite limited. They needed users to tag a dataset to train the machine to identify information, but this required a significant upfront investment by departments. Also, there was no clarity between the vendors we worked with on what constituted a sufficient number of tags, some said 30 000 tags for each subject/topic the machine was being asked to identify, other said 500 per subject/topic.

[Slide 22- Artificial Intelligence and machine learning in Records Management and Archives 3]

The study proved to us that automation is no longer a choice, but a necessity. However, that does not mean that humans/archivists are irrelevant in this process and we need to disabuse ourselves of the fear that it will make archivists, records managers or others redundant. The longer we maintain this belief and disengage from this topic the more problematic it will be for us in the long-term

The challenge with automating appraisal and selection, along with the sensitivity review process brings with it some important questions we need to consider:

- How do you measure accuracy? What does ‘good enough’ look like? What are the risks? What is acceptable risk appetite?
- How can we determine what might be missing?
- How can be accountable for the decisions we make based on machine outputs? How do we equally hold the machines to account?
- How do we compensate for the change in the digital record over time? Re-tune the algorithm?

We need to remember that we are likely to be dealing with ‘black boxes’ or algorithms created by commercial suppliers, so we need to understand how these systems work by testing and testing on lots of them with many kinds of data. If we do not engage with these systems and understand what they are doing but automate the records appraisal and selection process then we seriously risk biasing the historical records and by proxy history, along with our collective memories.

[Slide 23-Artificial Intelligence and machine learning in Records Management and Archives 4]

Emerging technologies such as artificial intelligence bring with them new and unexplored ethical challenges, which really require the profession to begin revising ethical codes to be able to guide archivists in navigating these waters. The difficulty in this work is that there are very few legislated and/or legal decisions to guide the profession on the appropriate ethical course of action that should be taken. There have been moves in the computer science community and technology industry to begin bounding acceptable parameters for algorithmic accountability by holding corporations and businesses accountable for how their machines arrive at a result. This is particularly important if these algorithms are being used to make decisions regarding people's lives ex: whether they receive a benefit, job etc. Currently software licenses place the burden of 'appropriate use' onto to users. The Association of Computer Machinery (ACM) wrote a declaration two years ago to set out seven principles for accountable and transparent algorithms: Awareness; access and redress; accountability; explanation; data provenance; auditability; validation and testing (https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf) Others have followed suit such as the Montreal Declaration https://www.declarationmontreal-iaresponsable.com/ The European Union has added some limited principles around AI that allows citizens to ask for an explanation as to how an algorithm arrived at a result, if it directly affects them. EU Regulations and principles around AI: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Finally a partnership on AI was developed between Google, Microsoft, IBM and Facebook to promote artificial intelligence for social good, although I have reservations about the outputs of a partnership driven by the largest international technology companies. (https://www.partnershiponai.org/)

We need to start testing these technologies in our practice to better understand the ethical implications AI has and in order to develop ethical codes that can support the community in their deployment and use. We also need to work more closely with computer science, data science, statisticians and others that are concerned with algorithmic accountability and transparency to ensure our voices our heard as well as to protect the integrity and authenticity of the record.

[Slide 24- Artificial Intelligence and Machine Learning in Research]

We have discussed the creation and acquisition of AI and machine learning and now let us look briefly at the automation of research along with its impact on records accessibility and re-use. The community has two issues to consider:

- Impact of researchers trying to mine archival data
- Digitisation of historical data and information

More and more researchers are starting to use data mining techniques to parse through large volumes of digital data. Examples include researchers using tools like Google NGRAM to mine digitize books to trace things like stereotypes in literature. See: Susan Mason. 'Analysing Stereotypes Across Time Using Google Ngram Viewer' SAGE Research Methods Cases Part 2 (2018) doi:10.4135/9781526436245. Using these data mining techniques to evaluate data across large information holdings has enable researchers to arrive at interesting conclusions that would not have been possible without the use of a machine.

There also many other tools being used by researchers, sometimes bespoke, that they use to mine large amounts of data, and this is no longer the preserve of the sciences and is now becoming more prevalent in the humanities in areas such as digital humanities, linguistics etc.

[Slide 25-Artificial Intelligence and Machine Learning in Research 2]

Whilst it is interesting to see these new advances in research there are questions that archivists need to consider around about how much access we may wish to allow researchers even to public records and data. We need to be cognizant that data mining and machine learning tools breakdown siloes created by archival description (i.e. fonds, series, files) which is both an opportunity to make new research discoveries but equally a threat because it can reveal unknown connections that become sensitive or problematic by virtue of making that connection. There are discrete pieces of information which alone are innocuous and unproblematic, but once a connection is made, they can reveal sensitive information. Essentially these technologies can surface confidential information that was missed during sensitivity review. Also once the data is mined and put into a system outside the archives, what else can it can be combined to?

Let's not get tunnel vision with AI. There is a danger of focusing too much on the impact on our individual collections, but what about linked data? And the semantic web? What will this mean for archives and opening our collections? We need to think about this now before implementing the enabling coding in our archives.

[Slide 26- Artificial Intelligence and Machine Learning in Research 3]

Researchers and the general public want us to digitise, everything if they could have their way, but as we know we do not have the resources to digitise everything and make it available. That said, we need to be more careful about what we digitise and how, given the advent of AI. In my opinion the re-purposing and re-use of archival records and data has enormous value and I think we sacrificed much of digitisation and allowing companies to digitize archival records and data, in order that we can get a 'free' copy'. We must be savvier. Companies are beginning to realise the value of data held in historical records. Digitising them and applying OCR is a method for gaining access to large volumes of data to train algorithms. We need to start asking ourselves:

- Why is the digitisation free?
- Will this data be used to train an algorithm?
- What is the company's ethical stance? Will they re-sell the data? If they are bought out what happens to the data?
- What happens to the data once the digitisation is done?
- Will there be an impact on people's lives?

We need to be more conscious of the potential impact of digitisation, especially for records that are rich in personal data.

[Slide 27- Conclusion]

I'm afraid that my conclusion brings more questions than answers but I think they are important things to consider as we develop our skills and understanding on the uses of AI.

- Government Use of Artificial Intelligence:
    - What role does the archives and information communities have to play in this space? Do we have a role?
    - What skills do we have, or do we need if we have a role to play?
    - What is the ‘record’? How do we capture and preserve that record?
    - Who are our partners? How do we begin to work with them?

- Machine Learning and Artificial Intelligence in Archival Processes
    - What is accuracy? What risks are we willing to accept?
    - How can we ensure the accountability of the decision we make based on machine-learning and AI processes?

- Artificial Intelligence and Machine Learning in Research
    - How much access is too much when machines are involved?
    - What are the right questions to ask when private companies offer us free digitisation?
    - How do researchers want to use our records to carry out digital research?