Edmund Balnaves

# 1 Artificial Intelligence and Libraries: An Introduction

**Abstract:** Artificial Intelligence (AI) is the use of computer systems to achieve tasks that would normally require human interpretive intervention. AI is integrally associated with Big Data and the Internet of Things, with large data stores of text, images, and videos on an extensive and wide-ranging scale providing the learning base. While AI does not purport to replace human cognition, the use of AI is becoming pervasive throughout society. It is present in smart phones to understand speech, help navigation and improve usability of apps. It is used in internet searching, surveillance systems, social credit systems, classification, and annotation systems. AI is used in advertising to match customers to products, and in social media to ensure content is exposed in a tightly directed way. AI is increasingly available or already present in the systems used by libraries. This chapter provides the background and context for this book *New Horizons in Artificial Intelligence in Libraries* which focuses on the current uses of AI in libraries and explores ways to prepare for and optimise the use of AI in libraries.

**Keywords:** Artificial intelligence; Big data; Internet of things; Machine learning

## What is Artificial Intelligence?

Artificial Intelligence (AI) is the use of computer systems to achieve tasks that would normally require human interpretive intervention and includes:

- Interpreting images for place, facial or object recognition
- Interpreting audio for language recognition and translation
- Analysing data in depth and breadth to discern patterns
- Controlling and managing movement of robotics in a real physical environment, and
- Conversational dialogue management that recognises, interprets and responds appropriately.

The concept of AI dates back to the early stages of computers and the philosophical examination of the implications of emerging computer technology, and in particular to a workshop in 1956 at Dartmouth College (McCarthy et al. 1995) that included the early greats of computer science John McCarthy and Marvin Minsky. The initial emphasis on expert systems gave way to the emergence of deep learning models.

Open Access. ©2025 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.
https://doi.org/10.1515/9783111336435-002

The value of AI in classification when used with large image stores was explored in projects such as the ImageNet Large Scale Visual Recognition (Krizhevsky, Sutskever, and Hinton 2012). Core to advances in AI in the last decade has been leveraging large online resources coupled with large-scale AI computing architectures to build models that can give credible predictive insights.

Big Data and the Internet of Things can be combined to operate at scale and in the real world. The core elements of AI bring together Big Data, text and image pattern recognition, natural language processing (NLP), and importantly the underpinning hardware capability to crunch through large data sets. The computer capabilities of massive parallel processing needed for AI originated with their use in the graphics processing units (GPUs) that enabled the hyperreal games now available on the market. GPUs and field-programmable gate arrays (FPGAs) provided the grunt needed for very large-scale model development and exploration of inferencing engines needed for AI (VerWay 2022, 36). AI toolkits, many written in the Python programming language or C, distil a range of complex pattern recognition algorithm and data management tools into simple and elegant processing pipelines to develop and train patterns in a large body of data to create models and apply real life data against the models to discern matches. These toolkits have provided the core to a rapidly evolving ecosystem of software applications covering all aspects of national language processing, pattern recognition and image recognition.

None of the work on AI has achieved human cognition and progress to date has not yielded the moment of singularity where computers become self-aware or achieve a level of artificial consciousness (Floridi and Chiriatti 2022). Nevertheless, use of AI tools embedded in many systems is becoming pervasive and persuasive. AI apps have been present on smart phones for many years and take an increasingly important role in extending the utility of the device to understand speech, dynamically translate audio inputs, assist with navigation and improve usability of apps.

The addition of AI to society creates a very new socio-political dynamic. While a few writers in the past forecast significant impacts of computer science developments, socialist theorists of the 19th and 20th centuries did not in general anticipate the effect on both labour and the political process the influence that AI systems might have. Potentially a class-based society is being supplanted by an information-based society. In its very nature, the information society is the perfect context for the application of artificial intelligence systems. In social media AI can be used to ensure that content emerges in a tightly directed way, with the potential to shape social ideas by biassing the content from social media systems in a particular direction. The Social Credit System in China involving a record system which tracks and evaluates businesses, individuals and government institutions for trustworthiness and China's Skynet ubiquitous surveillance system intrinsically depend on the tool-

sets of AI to achieve normative social goals. AI raises the possibility of at-scale social and behavioural intervention by both governments and global commercial entities such as Facebook, Google and Microsoft.

In the context of libraries, Bourg (2017) raises the question: what does it mean to maximise our collections for humans and what does it mean to maximise them for machines and algorithms? The process of re-envisaging the nature of the library has been gathering pace since the first library applications in machine learning emerged (Coleman 2017; Cordell 2020; Johnson 2018) and is exemplified in the formation of the International Federation of Library Associations and Institutions (IFLA) Artificial Intelligence Special Interest Group in 2022.

Central to the discussion of AI in libraries is the algorithm used to train AI systems to provide useful responses. Algorithm is the term used to encompass the set of machine learning techniques that may be deployed to glean intelligence from data, and probably more accurately refers to the process of model creation. The difficulty with the term algorithm is that it lends itself to an anthropomorphic interpretation of AI that is not accurate. Machine learning uses a range of techniques to harness the speed of computers in image, textual, location and other data points to derive an interpretive model from the data. The model can then be used to suggest a range of things: a face matched to a name, an emotional state or reaction, a set of structured terms to describe a document or image, a location based on an image, and many other similar activities. The underlying dataset and algorithm used in AI can lead to bias in the results and can be particularly visible in the ways in which AI can skew search results based on biases in the underlying training data (Snow 2018).

In 2018 the use of AI for initial job candidate selection emerged with the Russian bot Vera (Palmer 2018). Recent innovations include ChatGPT from OpenAI. GPT stands for Generative Pretrained Transformer and constitutes a model that takes the foundation of large text corpus modelling and integrates it with a neural network to parse the query and structure a logical response. The first GPT model was made available in 2018. Continual updates have been made and the current generation of the GPT architecture impressively goes beyond national language responses to provide code-meaningful responses to programming questions and semantically meaningful conversations and question responses. The semantic interpretation implied by the transformer neural networks is a generational advance in the capabilities of AI and has potentially profound implications in many areas.

An extension of ChatGPT is DALL-E, an AI image generating neural network built on a twelve billion-parameter version of GPT-3. DALL-E was trained on a data set of text-image pairs and can generate images from user-submitted text prompts. ChatGPT and DALL-E were both developed by OpenAI. The company name Open AI implies but does not mean that its systems are open-source with content available for others to use freely. DALL-E can creatively respond with image content to

queries based on requested text descriptions. For example, the following query: "Give me a logo for a search and discovery system that incorporates artificial intelligence recommendations" yielded the creative result (Figure 1.1).

**Figure 1.1:** Logo Created by AI for a Search and Discovery System

Microsoft made headlines in 2020 replacing journalists with machine-generated story selection rather than human-curated story selection for the MSN home page (Waterson 2020) and highlighted the risks of job displacement with the use of AI. The further advances represented by ChatGPT take the potential risks to another order. AI is capable of sophisticated responses across many disciplines. The answers provided to queries are well crafted and explanatory, even when they contain elements of factual errors.

## Implications of Artificial Intelligence for Libraries

Public libraries have grown in importance as gateways to information resources in a world where internet access is a prerequisite for engagement in community and social endeavours. The library has both explanatory and access roles. The transition to electronic collections heralded by Breeding (2011) has seen dramatic changes in research-oriented libraries. The preference for digital content has become entrenched in research institutions and the availability of AI-enhanced discovery tools is transforming the discovery, use and reuse of research data. The full implications of ChatGPT for research and reference librarians are yet to be seen. Some librarians may fear the development of AI in the research context, but others will embrace the potential and establish new and innovative services to enhance services provided to researchers.

Library integrated automation or management systems might be viewed as the precursors to AI within libraries. They have fundamentally changed the activities conducted by librarians, particularly as far as cataloguing and lending functions are concerned. Cataloguers have become metadata experts and subject analysts.

Cataloguing and classification have become shared activities across the world optimising collaborative efforts to improve the effectiveness of libraries and yield higher quality and consistency in item descriptions, content analysis, and information resource discovery. The developments have come at the practical cost of jobs on the ground for some librarians who have transferred to alternative roles. For example, the traditional bibliographic cataloguer once the quintessential exemplar of the library profession, is now a rare species.

The cataloguing role has changed considerably facilitated by library automation and touches the lives of fewer members of the profession. Outsourcing of library functions related to acquisition and selection of library collections and inclusion of the purchase of records with information resources has also changed activities. AI will be added to this mix to automate classification recommendations While there may be concern over algorithmic bias in machine learning systems, the bias in historical traditional classification schemes may be no better. The first edition of the Dewey Decimal Classification scheme is replete with unconscious and conscious bias (Mai 2016).

There is enormous potential for machine learning to enhance the quality of the work of librarians and to enrich access to the content of library collections held locally or accessed remotely in ways that might otherwise be impossible. The role of the metadata librarian is alive and well as the adjudicator of semantic description and content along with enhanced roles for reference librarians who can interact in new ways with large data and information stores.

The momentum of AI innovation and adoption is perhaps beyond the capability of libraries to control or manage. A rearguard action to block AI in what may well be a new industrial revolution in the making is not likely to be fruitful. By engaging with AI technology and enhancing knowledge and expertise concerning AI, the library and information profession can come to grips with its limits, risks, and opportunities. Through understanding the limits and capabilities of AI systems, librarians will be able to influence developments and provide new services to enhance information access and use for their communities.

There are many opportunities for the application of AI in libraries with the potential for the extension of activities in ways not currently possible. For instance, 24 x 7 assistance service can be provided through chatbots or in-library robotic assistants for check-in and checkout functions. The improvement of discovery for large collections of resources through automated classification, organisation, and retrieval makes possible at scale access to wonderful resources worldwide.

The debate regarding algorithmic bias is relevant to the library profession and the best way to engage in the debate is to participate in the work in the area. Many libraries are already participating in the AI debates. The bias in machine learning design can come from the source data sets which might be ethnically or culturally

homogeneous, or in machine learning rules or algorithms used. In social media and online news, the rules around story selection become all the more important where systematic bias may be an inadvertent or deliberate element in framing the information landscape presented to the user. AI systems may also have value in detecting and removing or rebalancing bias present in existing systems. AI systems could, for instance, assist with daunting task of correcting historical classification bias in historical bibliographic databases.

## Governance of Artificial Intelligence

There has been considerable debate within the community about the control of the use of AI and its future governance both within institutions and internationally. Institutional governance rules may increasingly incorporate mandates around the use of AI in the institution. The library profession can play a role in encouraging transparency in the use of AI. The debate around algorithmic bias is a case in point. The debates itself can lead to an awareness of the concerns and the need for adaptation and change in legislation around the world. The moves within the European Union to provide legislation related to AI (European Parliament 2023) and to enhance privacy and choice on the internet are cases in point. Algorithmic disclosure is all about transparency and the ability to understand the rules and assumptions that feed decision making. Openness enables a health public discourse about algorithms and their use and application.

The effectiveness of the open-source community illustrates the value of publishing coding and programming for scrutiny (Balnaves 2008). While openness permits attackers to understand vulnerability, it also facilitates discovery and resolution of vulnerabilities. Visibility fits well with the criteria for explainable machine learning outlined by Psychoula and others:

> There are several important elements to consider, and be able to explain, when creating automated decision-making systems:
> (1) The rationale behind the decision should be easily accessible and understandable
> (2) The system should maximise the accuracy and reliability of the decision.
> (3) The underlying data and features that may lead to bias or unfair decisions
> (4) The context in which the AI is deployed and the impact the automated decision might have for an individual or society (Psychoula et al. 2021, 50)

While the designers of AI systems might, for proprietary reasons, look askance at visibility into their systems, it is arguable that an open approach promotes a virtual cycle of improvement. Lack of regard for these principles was precisely the failure exhibited in implementing the Robodebt scheme in Australia (Shah 2023). The gov-

ernment took an automated algorithmic approach to assessing social security debt using model-based approximations. The result was the creation of debt notices based on estimates rather than actual social security debt figures, and in some cases the results were wildly wrong. The errors made resulted in great trauma for the people involved, including several deaths by suicide. The whole approach was ultimately found to be illegal in a subsequent class action brought against the government and harrowing insights from a Royal Commission (Australia. Royal Commission into the Robodebt Scheme 2023). Appropriate levels of governance and development frameworks along with legislative control are acutely important where AI is making critical judgements and will become even more necessary as AI becomes indistinguishable from a human response.

AI systems, once implemented, have potential risks, including spoofing or poisoning, where falsified data is fed to the system to skew the AI models developed (Laplante and Amaba 2021, 20). Where judgements are being made on an algorithmic basis, the risk of false positives is also a significant factor. Mortati and Carmel explore the issue of false positives in the case of cheating detection systems, an issue made all the more complex with the recent advent of ChatGPT (Mortati and Carmel 2021, 92).

How can libraries engage in AI firsthand? Chatbots are an uncomplicated way in which librarians can engage in the machine learning process to gain exposure to both the strengths and limitations of AI. Many tools for chatbot development are readily available online and allow free use. Most include natural language processing capabilities and methods to integrate additional textual analysis tools. Even without deploying chatbots, experimenting with them is an excellent way to understand both the opportunities and limitations of AI. Most people have undoubtedly experienced the friendly but hopelessly incompetent chatbots thrown up on websites by banks and other corporate entities as replacements for human interaction. They are examples of AI systems in their infancy. There is a long way to go in this area, but they will become ubiquitous in online systems.

Many toolkits for AI are open-source. For the adventurous who are willing to venture into coding, there are guides and examples abounding on the web to assist in experimenting with recommender systems which could make suggestions to library users.

The challenges of appropriate deployment of AI combine technological and governance elements. It is typical of any technological advance that governance and control elements lag behind the rapid advances occurring. The International Federation of Library Associations and Institutions (IFLA) has released a statement on libraries and AI that provides a good framework for considering the issues for AI in libraries (IFLA 2020). The European Union's High Level Expert Group on Arti-

ficial Intelligence has defined concerns in relation to AI and developed ethics guidelines which lists seven key requirements for trustworthiness of AI systems:

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and data governance
4. Transparency
5. Diversity, non-discrimination, and fairness
6. Societal and environmental wellbeing
7. Accountability (European Commission 2019, 14)

Librarians can engage with the vendors of their library systems to open discussion around the ways in which AI is used in their systems. There are many systems used by libraries where AI has the potential to add value. Each has different ethical dimensions some of which are explored by Jaillant and Caputo in the context of digital libraries. They argue:

> Instead, a framework of AI governance informed by well-developed language and procedures of consent, power, inclusivity, transparency, and ethics and privacy, inspired by consolidated practices among archivists as well in social science, historians and anthropologists, should drive the adoption of AI. (Jaillant and Caputo 2022, 833).

There is a *realpolitik* in institutions around investment that may make taking new directions difficult: an investment in major change like the implementation of AI may require a proof of concept initially to provide to the institution the value of the expenditure. Furthermore, the adoption of AI may come organically through the inclusion of AI in the software used in different sections of the library. The level of governance needs to be measured against the level of risk. More scrutiny is required where decisions about people and their livelihoods are involved, and less scrutiny perhaps where the AI implementation is targeted at an activity like reshelving books.

Many aspects of AI are beyond the governance of the library. The library's clients may be using agents such as ChatGPT and this use would be beyond the reach of the library's approach to governance as would the AI instruments built into the architecture of systems purchased for use by the library. The laws around use of AI are only now emerging, and lag in development and adoption behind implementation of AI. There is indeed a tension in national competition to stay on top of advances in AI, and the regulatory and governance elements that are emerging to protect against the potential and actual harms arising from use of AI. The responses at the national level vary according to the importance accorded to AI

and its governance by particular countries and vary significantly in relation to the democratic orientation of the legislature.

Governance issues are not the only ones that come into play with AI. The design of AI is grounded in models created from exceptionally large data sources of images, text, and video. The early efforts in natural language processing chatbots by Microsoft demonstrated the problems that analysis of such deep resources can present. The architects of large-scale AI consequently invest enormous funds into data sanitisation (Oprea, Singhal, and Vassilev 2022). The source datasets used may be skewed toward a particular group or area.

## Engagement by Libraries with Artificial Intelligence

Short of actions like Butlerian Jihad against computers found in Frank Herbert's science fiction novel *Dune*, there is little doubt that AI will become ever more entrenched in all systems. The choice is not rejection but engagement with the changes that will be wrought by the use of AI. The IFLA Information Technology Section has sponsored the creation of an Artificial Intelligence Special Interest Group within IFLA to foster debate and engagement around AI. The IT Section has held a series of conferences in Europe, Africa and Asia and South America as a means of strengthening the discussion on AI in the library field.

Going forward, there are many avenues for libraries to engage in AI. They include formulation of governance and ethical policies within libraries and any parent institutions. Activities taken by libraries may entail evaluation and scrutiny of library software that includes AI. Libraries can engage directly in the use of AI software through the many commercial and open-source tools now available. Libraries will be on the frontline for those asking about the implications of AI and will need to be well placed to assist with answers.

## References

Australia. Royal Commission into the Robodebt Scheme. 2023. *Report*. Canberra, ACT. https://robodebt.royalcommission.gov.au/system/files/2023-09/rrc-accessible-full-report.PDF.

Balnaves, Edmund. 2008."Open Source Library Management Systems: A Multidimensional Evaluation." *Australian Academic & Research Libraries* 39, no. 1: 1–13. https://doi.org/10.1080/00048623.2008.10721320. Available at https://www.academia.edu/99739844/Open_Source_Library_Management_Systems_A_Multidimensional_Evaluation?f_ri=7803.

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

Breeding, Marshall. 2011. “Preparing for the Long-Term Digital Future of Libraries,” *Computers in Libraries* 31, no. 1 (Jan./Feb.): 24–26. https://librarytechnology.org/document/15595.

Coleman, Catherine Nicole. 2017. “Artificial Intelligence and the Library of the Future, Revisited.” *Coffeehouse: A Collection of Data Blog Posts from Around the Web*, November 3, 2017. https://coffeehouse.dataone.org/2017/11/03/artificial-intelligence-and-the-library-of-the-future-revisted/.

Cordell, Ryan. 2020. *Machine Learning + Libraries: A Report on the State of the Field.* Commissioned by LC Labs, Library of Congress. Version published July 14, 2020. https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf.

European Commission. 2019. *Ethics Guidelines for Trustworthy AI*. Prepared by the High-Level Expert Group on AI. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html.

European Parliament. 2023. “EU AI Act: First Regulation on Artificial Intelligence.” *News European Parliament,* June 14, 2023. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

Floridi, Luciano, and Massimo Chiriatti. 2020. “GPT-3: Its Nature, Scope, Limits, and Consequences.” *Minds and Machines* 30, no. 4 (December): 681–94. https://doi.org/10.1007/s11023-020-09548-1.

International Federation of Library Associations and Institutions (IFLA). 2020. *IFLA Statement on Libraries and Artificial Intelligence.* [Prepared by] Committee on Freedom of Access to Information and Freedom of Expression (FAIFE). https://repository.ifla.org/handle/123456789/1646.

Jaillant, Lise, and Annalina Caputo. 2022. “Unlocking Digital Archives: Cross-Disciplinary Perspectives on AI and Born-Digital Data.” *AI & Society* 37, no. 3 (September): 823–35. https://doi.org/10.1007/s00146-021-01367-x.

Johnson, Ben. 2018. “Libraries in the Age of Artificial Intelligence.” *Computers in Libraries* 38, no.1 (January/February): 14–16. https://www.infotoday.com/cilmag/jan18/Johnson--Libraries-in-the-Age-of-Artificial-Intelligence.shtml#_blank.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2012. “ImageNet Classification with Deep Convolution Neural Networks.” *Advances in Neural Information Processing Systems* 25, 1106–1114. Available at https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

Laplante, Phil, and Ben Amaba. 2021. “Artificial Intelligence in Critical Infrastructure Systems” *Computer* 54, no. 10: 14–24. doi: 10.1109/MC.2021.3055892. Available at https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9548022.

McCarthy, John, M.L. Minsky, N. Rochester, and C.E. Shannon. 1955. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, Dartmouth College.” http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Mai, Jens-Erik. 2016. “Marginalization and Exclusion: Unraveling Systemic Bias in Classification.” *Knowledge Organization* 43, no. 5; 324–330. Available at https://www.nomos-elibrary.de/10.5771/0943-7444-2016-5-324.pdf .

Mortati, Joseph, and Erran Carmel. 2021. “Can We Prevent a Technological Arms Race in University Student Cheating?” *Computer* 54, no. 10: 90–94. doi: 10.1109/MC.2021.3099043. Available at https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9548108.

Oprea, Alina, Anoop Singhal, and Apostol Vassilev. 2022. “Poisoning Attacks Against Machine Learning: Can Machine Learning Be Trustworthy?” *Computer* 55, no. 11 (November): 94–99.

doi: 10.1109/MC.2022.3190787. Available at https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9928202.

Palmer, Annie. 2018. "Meet 'Vera,' the Russian AI that Could Vet You for Your Next Job: Pepsi, Ikea, and Other Major Firms Turn to Bizarre Humanoid to Find Candidates and Conduct Interviews." *Daily Mail. Australia,* March 30, 2018. https://www.dailymail.co.uk/sciencetech/article-5560547/Russian-AI-Vera-helping-companies-candidates-conduct-interviews.html.

Psychoula, Ismini, Andreas Gutmann, Pradip Mainali, Sharon H. Lee, Paul Dunphy, and Fabian A.P. Petitcolas. 2021. "Explainable Machine Learning for Fraud Detection." *Computer* 54, no. 10 (October): 49–59. doi: 10.1109/MC.2021.3081249. Available at https://arxiv.org/pdf/2105.06314.pdf?trk=public_post_comment-text.

Shah, Chiraag. 2023. "Australia's Robodebt Scheme: A Tragic Case of Public Policy Failure." *Voices: University of Oxford Blavatnik School of Government [Blog],* July 26, 2023. https://www.bsg.ox.ac.uk/blog/australias-robodebt-scheme-tragic-case-public-policy-failure#:~:text=Introduced%20in%202015%20as%20part,supposed%20overpayments%20to%20welfare%20recipients.

Snow, Jackie. 2018. "Bias Already Exists in Search Engine Results and It's Only Going to Get Worse." *MIT Technology Review*, February 26, 2018. https://www.technologyreview.com/2018/02/26/3299/meet-the-woman-who-searches-out-search-engines-bias-against-women-and-minorities/.

VerWey, John. 2022. "The Other Artificial Intelligence Hardware Problem." *Computer* 55, no. 1 (January): 34–42. https://doi.org/10.1109/mc.2021.3113271. Available at https://www.osti.gov/pages/servlets/purl/1847077.

Waterson, Jim. 2020. "Microsoft Sacks Journalists to Replace Them with Robots." *The Guardian*, May 30, 2020. https://www.theguardian.com/technology/2020/may/30/microsoft-sacks-journalists-to-replace-them-with-robots.