OLA IFC Tuesday Topics June 2021: Artificial Intelligence and Libraries

Welcome to Tuesday Topics, a monthly series covering topics with intellectual freedom implications for libraries of all types. Each message is prepared by a member of OLA's Intellectual Freedom Committee or a guest writer. Questions can be directed to the author of the topic or to the IFC Committee.

What is AI?

Artificial intelligence (AI) is no longer just a science fiction trope. In fact, AI technologies have become so prevalent in our lives over the past few years, that we encounter these technologies daily in applications like GPS navigation, online shopping recommendations, targeted ads, chatbots, virtual assistants, and search engines, to name just a few. Artificial Intelligence refers to all forms of machine learning, including deep neural networks, as well as computer vision, natural language processing, and other complex algorithms that attempt to replicate human decision-making. These technologies have a wide variety of applications, revealing patterns in data that would take human analysts eons to process. The American Library Association's Center for the Future of Libraries has identified Artificial Intelligence, along with several technologies powered by AI, such as facial recognition and self-driving cars, as top technology trends relevant to libraries. As libraries adopt these technologies, for things like digital text and image processing, algorithmic recommendations and discovery, and even virtual reference, it is important that we consider how this might impact our intellectual freedoms.

What are the IF concerns posed by AI?

As artificial intelligence transforms our lives and work, it has become clear that, as useful as it is, the technology poses numerous ethical concerns around bias, privacy, and misinformation. AI systems are trained on massive quantities of data, often harvested from social media and other online sources. There are concerns around how this data is collected, who controls it, how representative it is, and how it is used to target, profile, and manipulate. Libraries must contend with these issues as they implement AI tools in their own practices and as they help library users navigate digital life.

Digital privacy and consent

Most people have probably heard by now "if it's free online, you're the product," or more specifically, your data is the product. Whenever you click through a privacy agreement to use an app, you are most likely signing off on the collection, use, storage, and sale of your data. This includes your personal information, demographic data, location, and any and all interactions you have on the site. Big data is big business, and privacy policies are notoriously long and difficult to parse. As IFC member, Miranda Doyle, discussed in a previous Tuesday Topic about student privacy, libraries and schools should be safeguarding the privacy of their users when contracting with third party vendors. Library Freedom Project's scorecard rates the privacy practices of some of the most popular library vendors, providing a starting point for selecting and negotiating with vendors. Governments have also begun enacting regulations that give people more control over what data they share. However, many people don't realize that anything on the web is easily scraped by outside companies,

researchers, or individuals that want to harvest data, no consent required. Artificial intelligence developers frequently purchase or scrape the large quantities of data they need for machine learning projects. Such developers include companies like Clearview A.I., which secretly developed a real-time facial recognition database from images and data scraped from the open web.

Replicating and reinforcing bias

MIT researcher Joy Buolamwini's work, as shown in the film Coded Bias, has drawn attention to the fact that when facial recognition systems are trained on mostly white male faces, they perform poorly at identifying non-white or non-male faces, often misidentifying or failing to identify them. This has led to wrongful arrests, like the case of Robert Williams in Detroit, and has spurred dozens of cities, including Portland, to ban the use of facial recognition technology by law enforcement and public agencies. Although this issue is most visible with facial recognition, the truth is that any machine learning system replicates, reinforces, and sometimes amplifies biases in the data it is trained on. The "black box" nature of AI algorithms can lead people to believe that the decisions they make are fairer than those made by humans, but as anyone who works with data knows, "garbage in, garbage out," and much of the data fed into machine learning systems is not cleansed of the racist, sexist, and classist garbage of the society that produced it. Virginia Eubanks' book *Automating Inequality* and Cathy O'Neill's *Weapons of Math Destruction* both detail the harms caused by black box algorithms, when they are unchecked by human empathy and judgment. This problem of bias in AI systems has been recognized in fields as wide ranging as medical diagnostics, predictive policing, and even search engines. Safiya Noble's book, *Algorithms of Oppression* is a seminal work on understanding the bias implicit in the search algorithms we rely on every day and the ways they profile and misrepresent BIPOC and other marginalized demographic groups.

Deep fakes and viral misinformation

Another threat to intellectual freedom is the rise of misinformation produced and spread by artificial intelligence systems. Artificially produced content, including text, images, and video, has become so sophisticated, that it is nearly indistinguishable from real content. Malicious actors have used fake AI-produced content to interfere with elections and sow widespread confusion. Social media algorithms that prioritize high levels of engagement, have also been shown to spread misinformation more quickly than true stories. Ironically, AI has also been proposed as a solution to the problem of misinformation, as the same tools that are able to produce fake content are best able to detect it. However, this raises questions about how much trust to put in such moderation algorithms to define what is true.

What can libraries do?

AI is here to stay. It offers undeniable benefits, such as better accessibility with speech to text and conversational searching, tools for managing and analyzing digital documents, and improved search and discovery. Libraries are experimenting with AI to improve optical character recognition in text documents, automate processing of digital images, provide recommendations to users based on their past searches, and for virtual reference assistance. These uses of AI are potentially transformative, but the ethical issues inherent in these technologies also threaten values that librarians hold dear. Libraries using AI applications need to be aware of these intellectual freedom issues, work to mitigate them, and help users to understand them.

As digital literacy advocates and conduits to emerging technologies, libraries have an opportunity to demystify and democratize AI technologies and advocate for more equitable data practices. In 2019 the Urban Libraries Council launched an AI and Digital Citizenship initiative, calling for libraries to get ahead of the curve by educating ourselves and our users about AI and incorporating the technology into our services in an ethical and transparent way. The International Federation of Library Associations (IFLA) also advocates for ethical use

of AI in libraries. Libraries have partnered with AI researchers to develop library specific apps and programs to teach people about AI and with advocacy organizations to provide creative programs about issues of bias and surveillance. Libraries have also enabled hands-on exploration through maker kits or lab spaces and encouraged civic engagement by hosting community conversations. Providing broad access to AI technology and helping people understand it is a first step toward diversifying the artificial intelligence workforce, developing public policy solutions, and reducing the problem of bias. Informed librarians can also provide human guidance, such as inclusive data curation and technological literacy, to mitigate the harms of unchecked algorithms.

Ellie Avis
OLA Intellectual Freedom Committee Member
Technical Services Manager, Josephine Community Library

<p align="center">Learn More:</p>

AI in Libraries

Center for the Future of Libraries. (n.d.). *Trends*. American Library Association.
http://www.ala.org/tools/future/trends

Finley, T. K. (2019). The Democratization of Artificial Intelligence: One Library's Approach. *Information Technology and Libraries*, *38*(1), 8-13. https://doi.org/10.6017/ital.v38i1.10974

Garcia-Febo, L. (2019, March 1). Exploring AI: How libraries are starting to apply artificial intelligence in their work . *American Libraries*. https://americanlibrariesmagazine.org/2019/03/01/exploring-ai/

Ghosh, S. (2021, March 15). *Future of AI in libraries*. SJSU School of Information.
https://ischool.sjsu.edu/ciri-blog/future-ai-libraries

IFLA. (2020, October 21). *Statement on Libraries and Artificial Intelligence.*
https://www.ifla.org/publications/node/93397

Wheatley, A & Hervieux, S. (2020). Artificial intelligence in academic libraries: An environmental scan. *Information Services & Use, 39*(4), 347–356.

Algorithmic Bias

Altman, A. (2021, May 20). Users, bias, and sustainability in A.I. *Digital Public Library of America.*
https://dp.la/news/users-bias-and-sustainability-in-ai

Coded Bias. (2020). https://www.codedbias.com

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

Gilman, M. & Madden, M. (2021). Digital barriers to economic justice in the wake of COVID-19. *Data & Society.*
https://datasociety.net/library/digital-barriers-to-economic-justice-in-the-wake-of-covid-19

Harwell, D. (2021, April 13). Wrongfully arrested man sues Detroit police over false facial recognition match. *Washington Post.*
https://www.washingtonpost.com/technology/2021/04/13/facial-recognition-false-arrest-lawsuit/

Noble, S. U. (2019). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neill, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy.* Crown.

Snow, J. (2018, July 26). Amazon's face recognition falsely matched 28 members of Congress with mugshots. *American Civil Liberties Union.* https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

Metz, C. (2021, March 15). Who is making sure the A.I. machines aren't racist? *New York Times*. https://www.nytimes.com/2021/03/15/technology/artificial-intelligence-google-bias.html

Wiggers, K. (2021, February 18). Studies find bias in AI models that recommend treatments and diagnose diseases. *Venture Beat.* https://venturebeat.com/2021/02/18/studies-find-racial-and-gender-bias-in-ai-models-that-recommend-ventilator-usage-and-diagnose-diseases/

Laws and Regulations

Keane, J. (2021, April 8). From California to Brazil: Europe's privacy laws have created a recipe for the world. *CNBC*. https://www.cnbc.com/2021/04/08/from-california-to-brazil-gdpr-has-created-recipe-for-the-world.html

Metz, R. (2020, September 9). Portland passes broadest facial recognition ban in the US. *CNN Business.* https://www.cnn.com/2020/09/09/tech/portland-facial-recognition-ban/index.html

Misinformation

Kreps, S. (2020, June). *The role of technology in online misinformation*. Foreign Policy at Brookings. https://www.brookings.edu/wp-content/uploads/2020/06/The-role-of-technology-in-online-misinformation.pdf

Meserole, C. (2018, May 9). How misinformation spreads on social media—And what to do about it. *Order from Chaos.* Brookings. https://www.brookings.edu/blog/order-from-chaos/2018/05/09/how-misinformation-spreads-on-social-media-and-what-to-do-about-it/

Digital Privacy

Hill, K. (2021, March 18). Your face is not your own. *New York Times Magazine.* https://www.nytimes.com/interactive/2021/03/18/magazine/facial-recognition-clearview-ai.html

Hodge, K. (2018, April 19). If it's free online, you are the product. *The Conversation.* https://theconversation.com/if-its-free-online-you-are-the-product-95182

Roderick, M. (2021, January 18). Visualizing the length of privacy policies. *USDirect.* *https://www.usdirect.com/business/resource-center/privacy-policy-lengths/*