**PROJECT**
Information Literacy

## Further Reading | The Age of Algorithms

## About Algorithms

Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner (15 November 2019), "Americans and privacy: Concerned, confused, and feeling a lack of control over their personal information," Pew Research Center. https://pewrsr.ch/37Fdad3

> "A majority of Americans believe their online and offline activities are being tracked and monitored by companies and the government with some regularity. It is such a common condition of modern life that roughly six-in-ten U.S. adults say they do not think it is possible to go through daily life *without having data collected about them* by companies or the government."

Aaron Smith (16 November 2018), "Public attitudes toward computer algorithms," Pew Research Center. https://pewrsr.ch/2ZRAOjG

> "When it comes to the algorithms that underpin the social media environment, users' comfort level with sharing their personal information also depends heavily on how and why their data are being used. A 75% majority of social media users say they would be comfortable sharing their data with those sites if it were used to recommend events they might like to attend. But that share falls to just 37% if their data are being used to deliver messages from political campaigns."

Anjana Susarla (17 April 2019), "The new digital divide is between people who opt out of algorithms and people who don't," *The Conversation.* https://bit.ly/2ZQr49y

> "... as digital devices proliferate, the divide is no longer just about access. How do people deal with information overload and the plethora of algorithmic decisions that permeate every aspect of their lives? The savvier users are navigating away from devices and becoming aware about how algorithms affect their lives. Meanwhile, consumers who have less information are relying even more on algorithms to guide their decisions."

Zeynep Tufekci (September 2017), "We're building a dystopia just to make people click on ads," TED video. https://bit.ly/2T6fCp5

> "We're building an artificial intelligence-powered dystopia, one click at a time, says techno-sociologist Zeynep Tufekci. In an eye-opening talk, she details how the same algorithms companies like Facebook, Google and Amazon use to get you to click on ads are also used to organize your access to political and social information. And the machines aren't even the real threat. What we need to understand is how the powerful might use AI to control us — and what we can do in response." See also her 2017 book *Twitter and Teargas: The Power and Fragility of Networked Protest.*

Jonathan Zittrain (23 July 2019), "The hidden costs of automated thinking," *The New Yorker*. https://bit.ly/39IthZh

> "In the past, intellectual debt has been confined to a few areas amenable to trial-and-error discovery, such as medicine. But that may be changing, as new techniques in artificial intelligence — specifically, machine learning — increase our collective intellectual credit line. Machine-learning systems work by identifying patterns in oceans of data. Using those patterns, they hazard answers to fuzzy, open-ended questions. Provide a neural network with labelled pictures of cats and other, non-feline objects, and it will learn to distinguish cats from everything else; give it access to medical records, and it can attempt to predict a new hospital patient's likelihood of dying. And yet, most machine-learning systems don't uncover causal mechanisms. They are statistical-correlation engines. They can't explain why they think some patients are more likely to die, because they don't "think" in any colloquial sense of the word — they only answer. As we begin to integrate their insights into our lives, we will, collectively, begin to rack up more and more intellectual debt."

## Algorithms, Social Justice, and Implications for Personal Agency

Amnesty International (November 2019) *Surveillance giants: How the business models of Google and Facebook threaten human rights.* https://bit.ly/2um0Ik6

> "These algorithmic systems have been shown to have a range of knock-on effects that pose a serious threat to people's rights, including freedom of expression and opinion, freedom of thought, and the right to equality and non-discrimination. These risks are greatly heightened by the size and reach of Google and Facebook's platforms, enabling human rights harm at a population scale. Moreover, systems that rely on complex data analytics can be opaque even to computer scientists, let alone the billions of people whose data is being processed."

Ruha Benjamin (25 October 2019), "Assessing risk, automating racism," *Science.* https://bit.ly/2QqMveb

> "Jim Crow practices feed the "New Jim Code" — automated systems that hide, speed, and deepen racial discrimination behind a veneer of technical neutrality. Data used to train automated systems are typically historic and, in the context of health care, this history entails segregated hospital facilities, racist medical curricula, and unequal insurance structures, among other factors. Yet many industries and organizations well beyond health care are incorporating automated tools, from education and banking to policing and housing, with the promise that algorithmic decisions are less biased than their human counterpart. But human decisions comprise the data and shape the design of algorithms, now hidden by the promise of neutrality and with the power to unjustly discriminate at a much larger scale than biased individuals." See also her 2019 book, *Race after technology: Abolitionist tools for the new Jim Code.*

Jessie Daniels (19 October 2017), "Twitter and white supremacy, a love story," *DAME.* https://bit.ly/2FpLEnU

> "When Twitter launched in 2006, it unwittingly gave white supremacists an ideal venue for their hatred. Social media experts like to talk about the "design affordances" of a platform, meaning the built-in clues that suggest how a platform is meant to be used. Twitter gained a reputation among some users for its use of hashtags for breaking news and for organizing, as in the Arab Spring in 2010 and Black Lives Matter in 2013.  For ideologically committed white supremacists, the affordances of Twitter pointed to new mechanisms for the furtive spread of propaganda and for vicious harassment with little accountability. The rise of social media platforms like Twitter, 4chan, and Reddit, meant

that white nationalists had many places to go online besides Stormfront. It also meant that the spread of white nationalist symbols and ideas could be accelerated and amplified by algorithms." See also her prescient 2009 book, *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights.*

Safiya Noble (26 March 2018), "Google has a striking history of bias against black girls" *Time.* https://bit.ly/36r8DuF

"My first encounter with racism in search was in 2009 when I was talking to a friend who casually mentioned one day, 'You should see what happens when you Google 'black girls.' I did and was stunned . . .I encourage us all to take notice and to reconsider the affordances and the consequences of our hyper-reliance on these technologies as they shift and take on more import over time. What we need now, more than ever, is public policy that advocates protections from the effects of unregulated and unethical artificial intelligence." See also her 2018 book, *Algorithms of oppression: How search engines reinforce racism*.

Mark McCarthy (15 March 2019), "The ethical character of algorithms," Shorenstein Center, Kennedy School of Government, Harvard University. https://bit.ly/37vOxPY

"Now the algorithm is king. Algorithms are increasingly used for consequential decision-making in all areas of life. The same questions that troubled these earlier scholars arise again with renewed urgency. Are these mathematical formulas expressed in computer programs value-free tools that can give us an accurate picture of social reality upon which to base our decisions? Or are they intrinsically ethical in character, unavoidably embodying political and normative considerations? Do algorithms have politics, or does it all depend on how they are used?" Includes analysis of algorithmic personalization in journalism.

Siva Vaidhyanathan (25 November 2019), "Digital democracy will face its greatest test in 2020," *The Guardian*. https://bit.ly/2QPuW6p

"The shocks of 2016 awakened journalists and regulators to the ways that social media undermines democracy. After a decade of shallow proclamations of their democratic potential, it's clear that Facebook, Twitter and Google are, in fact, major threats to democracy . . . we should attend to the places where citizens have little but Facebook through which to view their countries, governments and the world." See also his 2018 book, *Antisocial Media: How Facebook disconnects us and undermines democracy.*

## Algorithms, Higher Education, and Libraries

Drew Harwell (24 December 2019), "Colleges are turning students' phones into surveillance machines, tracking the locations of hundreds of thousands," *The Washington Post.* https://wapo.st/37Ajfrd

"If a tracking system can make students be better, one college adviser said, isn't that a good thing? But the perils of increasingly intimate supervision — and the subtle way it can mold how people act — have also led some to worry whether anyone will truly know when all this surveillance has gone too far. 'Graduates will be well prepared to embrace 24/7 government tracking and social credit systems,' one commenter on the Slashdot message board said. 'Building technology was a lot more fun before it went all 1984.'"

Kyle M. L. Jones et al (April 2019), "In their own words: Student perspectives on privacy and library participation in learning analytics initiatives," *Recasting the Narrative: Proceedings of the Association of College and Research Libraries Conference,* April 10-13, 2019. Cleveland, OH. https://bit.ly/35qJVZZ

> "Students are generally unaware of the data and information their institutions have access to about themselves. Interviews prompted students to list examples of student information, but this proved difficult. As one student said, 'I don't know what information [my institution] is necessarily taking from me.' Probing questions elicited responses indicating that students expect their institution to record demographic information (e.g., names, addresses, and phone numbers), financial aid information, and academic information, such as the courses in which they enrolled and the grades they earned. Once students began identifying types of information to which their institution had access, they would also begin exploring information sources. Often, students recognized that a learning management system was an environment that could capture, as one participant said, 'every move a student is making.' Other students recognized that using their student ID card or connecting to campus WiFi may produce data as well."

Annemaree Lloyd (2019), "Chasing Frankenstein's monster: Information literacy in the black box society," *Journal of Documentation*, *75*(6), 1475-1485. https://bit.ly/2SW7JSI

> "To build a critically reflexive approach to algorithms into information literacy pedagogy, key concepts such as bias, trust, credibility, opacity, diversity, and social justice, commensurability (how algorithms interact with us to shape and reshape knowledge and agency) and performativity should be incorporated to supplement and deepen concepts such as search, and the core activities associated with current information literacy practice. In this respect, algorithmic literacy differs from digital literacy, which focuses on core information literacy skills in the digital context, because it requires examination of culture (in both analogue and digital spaces), as a generative proposition and the construction of algorithms should be viewed as a practice which influences other aspects of social life."

Matthew Reidsma (11 March 2016), "Algorithmic bias in library systems" (blog post). https://bit.ly/2MYDcjH

> More and more academic libraries have invested in discovery layers, the centralized "Google-like" search tool that returns results from different services and providers by searching a centralized index. The move to discovery has been driven by the ascendance of Google as well as libraries' increasing focus on user experience. Unlike the vendor-specific search tools or federated searches of the previous decade, discovery presents a simplified picture of the library research process. It has the familiar single search box, and the results are not broken out by provider or format but are all shown together in a list, aping the Google model for search results. Discovery's promise of a simple search experience works for users, more often than not. But discovery's external simplicity hides a complex system running in the background, making decisions for our users. And it is the rare user that questions these decisions. See also his 2019 book, *Masked by trust: Bias in library discovery*.

## Algorithms and Journalism

Pete Brown, Andrea Wenzel and Meritxell Roca-Sales (17 October 2017), "Local audiences consuming news on social platforms are hungry for transparency," *CJR Tow Center Reports*. https://bit.ly/39JXJls

> "At one end of the scale there was an observable lack of awareness about the existence and/or purpose of the algorithms that control the flow of news on their news feeds. At the other, where

people had some awareness of algorithms, we frequently observed a) a framing of them as "filters" that create one-sided filter bubbles or b) a perception that they, as autonomous individuals, wield more power over what news they see on platforms than the platforms' algorithms."

Nicholas Diakopoulos (28 November 2018), "An algorithmic nose for news," *Columbia Journalism Review.* https://bit.ly/2QWPNVE

Algorithmic claim spotting is one of a growing number of applications of computational story discovery. Whether monitoring political campaign donations, keeping an eye on the courts, surfacing newsworthy events based on social media posts, winnowing down hundreds of thousands of documents for an investigation, or identifying newsworthy patterns in large datasets, computational story discovery tools are helping to speed up and scale up journalists' ability to surveil the world for interesting news stories. Algorithms offer a sort of data-driven sixth sense that can help orient journalistic attention.

Bernat Ivancsics and Mark Hansen (21 November 2019),  "Actually, it's about ethics, AI, and journalism: Reporting on and with computation and data," *CJR Tow Center Reports.* https://bit.ly/36qvs1j

"So every reporting beat is now a data beat, and computation is an essential tool for investigation. But digitization is affected by inequities, leaving gaps that often reflect the very disparities reporters seek to illustrate. Computation is creating new systems of power and inequality in the world. We rely on journalists, the 'explainers of last resort', to hold these new constellations of power to account. We report *on* computation, not just *with* computation."

Jihii Jolly (20 May 2014), "How algorithms decide the news you see," *Columbia Journalism Review.* https://bit.ly/39Hixuc

"While publishers view optimizing sites for the reading and sharing preferences of specific online audiences as a good thing, because it gets users to content they are likely to care about quickly and efficiently, that kind of catering may not be good for readers."

Tim Libert and Reuben Binns (2019), "Good news for people who love bad news: Centralization, privacy, and transparency on US news sites," WebSci '19, June 30-July 3, 2019, Boston, MA. https://bit.ly/2FpNftU

In this study, 4,000 US-based news sites, 4,000 non-news sites,and privacy policies for 1,892 news sites and 2,194 non-news sites are examined. We find news sites are more reliant on third-parties than non-news sites, user privacy is compromised to a greater degree on news sites, and privacy policies lack transparency in regards to observed tracking behaviors. Overall, findings indicate the democratic role of the press is being undermined by reliance on the "surveillance capitalism" funding model.

Neil Thurman, Seth C. Lewis, and Jessica Kunert (2019), "Algorithms, automation, and news," *Digital Journalism 7*(8), 980-992. https://bit.ly/39QL6p0

Extensively referenced introduction to a theme issue of the journal. "By the mid-2010s, it had become clear that fully automated and semi-automated forms of gathering, filtering, composing, and sharing news had assumed a greater place in a growing number of newsrooms (Diakopoulos 2019; Dörr 2016), opening the possibility that there were places where shifts in the norms, patterns, and routines of news production were happening and even that, at a more fundamental level, taken-for-granted ideas about who (or what) does journalism were being challenged (Lewis, Guzman, and Schmidt 2019; Primo and Zago 2015)."

**Teaching Resources**

All Hail the Algorithm (short documentaries from Al Jazeera) https://bit.ly/35o0ohl
Check, Please! (lesson plans) http://lessons.checkplease.cc
Civic Online Reasoning (lesson plans) https://cor.stanford.edu/
Clickbait, Bias, and Propaganda in Information Networks (OER textbook) https://bit.ly/2MZ6HSm
Kaitlin L. Costello, Critical Algorithm Studies (syllabus) https://bit.ly/35rytgs
Data Detox Kit (resource toolkit) https://datadetoxkit.org/
Do Not Track - (personalized documentary series) https://donottrack-doc.com/
Fairness Toolkit (resource toolkit) https://unbias.wp.horizon.ac.uk/fairness-toolkit/
Ten weird tricks for resisting surveillance capitalism in and through the classroom (classroom ideas) https://bit.ly/2FpB93T
Andrea L. Guzman, AI, Automation and Journalism (journalism seminar syllabus) https://bit.ly/39HtjjQ
Surveillance Self-Defense (resource toolkit) https://ssd.eff.org/

**Organizations and Websites for Keeping Up**

ACLU Privacy and Technology News, https://www.aclu.org/news/by-issue/privacy-technology/
AINow Institute https://ainowinstitute.org/
The Algorithm (newsletter from MIT Technology Review)
https://forms.technologyreview.com/the-algorithm/
AlgorithmWatch https://algorithmwatch.org/en/
Berkman Klein Center for Internet and Societyat Harvard University https://cyber.harvard.edu/publications
Center for Democracy and Technology https://cdt.org/
CyLAb CMU Security and Privacy Institute https://www.cylab.cmu.edu/
Data & Society https://datasociety.net/output/
Data Doubles https://datadoubles.org/publications/
Datajournalism.com https://datajournalism.com/ and the newsletter Conversations with Data
https://datajournalism.com/read/newsletters
Digital Watch Observatory https://dig.watch/
Electronic Privacy Information Center (EPIC) https://epic.org/
EthicsLab https://ethicslab.georgetown.edu/
Harvard University Privacy Tools Project https://privacytools.seas.harvard.edu/
Oxford Internet Institute blog https://www.oii.ox.ac.uk/blog/
tactical tech https://tacticaltech.org/