



INFORMATION LITERACY IN THE AGE OF ALGORITHMS

Student experiences with news and information,
and the need for change

Alison J. Head, Ph.D.
Barbara Fister
Margy MacMillan



Contents

Introduction	1
Part One: The Age of Algorithms	3
Part Two: Student Perceptions and Concerns	13
Part Three: Conclusion and Recommendations	27
Leading Thinkers	37
Methods	43
Keywords and Definitions	49
About PIL	51

Introduction

We live in an era of ambient information. Amidst the daily flood of digital news, memes, opinion, advertising, and propaganda, there is rising concern about how popular platforms, and the algorithms they increasingly employ, may influence our lives, deepen divisions in society, and foment polarization, extremism, and distrust. For the past decade, Project Information Literacy (PIL)¹ has conducted large-scale studies on how college students interact with information for school, for life, for work, and most recently, for engaging with the news. The latest report from PIL stands at the intersection of these critical information practices and asks: How much do students know about how internet giants such as Google, YouTube, Instagram, and Facebook work, and, in turn, how they work on society?

This pivotal generation² born before the constant connectivity of social media, has come of age aware, cautious, and curious about the implications of the current information landscape. Deeply skeptical, many of these students are conditioned to do research for themselves rather than deferring to experts or major news outlets. They understand that “free” platforms are convenient but also recognize they harvest massive amounts of personal data to target ads and influence what content they see.

While many students worry about how the next generation will fare in terms of disinformation,

privacy, and personal well-being, they do not fully understand how big data and artificial intelligence (AI) are being used in educational technology and society. Neither do their professors. While faculty are alarmed about the social impact of the internet giants and the loss of a common culture, they have little idea how to relate their concerns to the information tasks intrinsic to the courses they teach.

When librarians and educators first adopted information literacy and critical thinking as essential educational outcomes, the algorithm-driven platforms many of us turn to — Google, YouTube, Facebook, Instagram, and Amazon — did not exist. Though information literacy has grown to include news literacy in the wake of the “fake news” crisis,³ there is little consideration of how colossal sites like these influence what we see and learn, what we think, and ultimately, who we are.

If we believe that information literacy educates students for life as free human beings who have the capacity to influence the world, then information literacy needs to incorporate an understanding of ways that news and information flows are shaped by algorithms. To do this, we need to know more about how students interact with algorithm-driven platforms. We must consider courses of action for educators preparing students to understand the technological and social forces shaping the circulation of news and information in society today.

1 For details about Project Information Literacy, see <https://www.projectinfoilit.org/about.html>

2 See “The pivot generation” on p. 26 for a discussion of how students in this study described themselves as belonging to a distinct age cohort because of their childhood experiences with emerging technologies.

3 Alison J. Head, Erica DeFrain, Barbara Fister, and Margy MacMillan (5 August 2019), “Across the great divide: How today’s college students engage with news,” *First Monday* 24(8), DOI: <https://doi.org/10.5210/fm.v24i8.10166>

In the growing research literature about students and algorithms, two recent studies help inform these efforts. A much-discussed 2018 survey of more than 4,500 Americans revealed widespread concerns about computer algorithms making automated decisions with real-life consequences, such as who gets a job or a loan.⁴ In 2017, a survey of college students found most were unaware of whether or not the news they got from Google and through Facebook was filtered using algorithms.⁵ Many questions remain, however, about what students already know, and need to know, about the individual and social effects of algorithmic filters.

Three sets of questions guided this report's inquiry:

1. What is the nature of our current information environment, and how has it influenced how we access, evaluate, and create knowledge today? What do findings from a decade of PIL research tell us about the information skills and habits students will need for the future?
2. How aware are current students of the algorithms that filter and shape the news and information they encounter daily? What concerns do they have about how automated decision-making systems may influence us, divide us, and deepen inequalities?
3. What must higher education do to prepare students to understand the new media landscape so they will be able to participate in sharing and creating information responsibly in a changing and challenged world?

To investigate these questions, we draw on qualitative data that PIL researchers collected from student focus groups and faculty interviews during fall 2019 at eight U.S. colleges and

universities. Findings from a sample of 103 students and 37 professors reveal levels of awareness and concerns about the age of algorithms on college campuses. They are presented as research takeaways.

This report is divided into three parts:

Part One describes what has been called “the age of algorithms,” and discusses catalysts for a profound shift in the information landscape, including big data, automated decision making, and AI. A critical review of PIL's ongoing research highlights what students, and the faculty who teach them, need to know.

Part Two presents findings from student focus groups and faculty interviews to explore campus-wide awareness about automated decision-making, personalization of information and news, and concerns about how this may influence students' knowing and learning.

Part Three makes four recommendations for stakeholders — educators, librarians, administrators, and journalists — that consider possibilities for reimagining information literacy in light of new and dramatically different circumstances. To explore the implications of this study's findings, we convened an interdisciplinary group of leading thinkers⁶ in education, libraries, media research, journalism, and technology to explore the challenges and opportunities we face. Concise commentaries from participants are featured at the end of this report.

This 10th anniversary report takes PIL into the age of algorithms. It summons educators and librarians to embrace the considerable challenge of understanding the technological and social forces shaping the circulation of news and information in

4 Aaron Smith (16 November 2018), “Public attitudes toward computer algorithms,” Pew Research Center, <https://www.pewresearch.org/internet/2018/11/16/public-attitudes-toward-computer-algorithms/>

5 The survey sample size was N = 147 college students, plus interviews with 37 faculty were also conducted as part of this study. Elia Powers (2017), “My news feed is filtered? Awareness of news personalization among college students,” *Digital Journalism* 5(10), 1315-1335, DOI: <https://doi.org/10.1080/21670811.2017.1286943>

6 The “Algo Study Thinking Leader Session” was held at Harvard Graduate School of Education on November 7, 2019.

contemporary society. It provides qualitative data that tell us how college students conceptualize and navigate a volatile and ever-changing information landscape. It explores what can be done to prepare students to deal with this new reality.

To date, no systematic investigation has explored what college students and faculty think about algorithm-driven platforms and concerns they may have about their privacy and access to trustworthy news and information.

Part One: The Age of Algorithms

Algorithms — rule-based processes for solving problems — predate computers. It was not, however, until the word “Google” became synonymous with “to search online” in the early 2000s⁷ that the idea of algorithms entered the public consciousness.⁸ That was when we began to notice how clever computer code influenced our daily lives by recommending Netflix films,⁹ remembering preferences for Amazon purchases, and finding our friends on Facebook’s precursors, such as Friendster and MySpace.

Within a few years of its founding in 1998, Google needed a profitable business model, so it began to make use of the digital trails we all leave behind to profit from personalized advertising. Facebook soon followed. The behemoth social media platform built its reputation and advertising might on its “social

graph,” the interconnections among people online, enriched by metrics of “friends” and “likes.”¹⁰

During the same time, the news industry began to struggle as startups like Craigslist began to cannibalize classified ad revenues and subscriptions dwindled as readers enjoyed free news online.¹¹ News organizations were forced to negotiate fraught relationships with platforms that increasingly dominated both digital advertising and monopolized audience attention.¹²

Fast forward to 2015, when controversies around “fake news” and the splintering of global audiences into polarized camps led the public to view algorithms as powerful, efficient — and often questionable — drivers of innovation and social change. The rise of what is widely known as the “age of algorithms” has had a profound impact

7 Virginia Heffernan (15 November 2017), “Just Google it: A short history of a newfound verb,” *Wired*, <https://www.wired.com/story/just-google-it-a-short-history-of-a-newfound-verb/>

8 Tarleton Gillespie (2017), “Algorithm,” *Digital keywords: A vocabulary of information society and culture*, Benjamin Peters (Ed.), Princeton University Press, 18-29.

9 Clive Thompson (21 November 2008), “If you liked this, you’re sure to love that,” *The New York Times Magazine*, <https://www.nytimes.com/2008/11/23/magazine/23Netflix-t.html>

10 Jonathan Haidt and Tobias Rose-Stockwell (December 2019), “The dark psychology of social networks,” *The Atlantic*, <https://www.theatlantic.com/magazine/archive/2019/12/social-media-democracy/600763/>

11 Ángel Arrese (2016), “From gratis to paywalls: A brief history of a retro-innovation in the press’s business,” *Journalism Studies* 17(8), 1051–1067, DOI: <https://doi.org/10.1080/1461670X.2015.1027788>

12 Nushin Rashidian, George Tsiveriotis, and Pete Brown — with Emily Bell and Abigail Hartstone (22 November 2019), *Platforms and publishers: The end of an era*, Tow Center for Digital Journalism, https://www.cjr.org/tow_center_reports/platforms-and-publishers-end-of-an-era.php

on society,¹³ on politics,¹⁴ on the news,¹⁵ and on epistemology.^{16,17} And yet, most algorithms are easy to ignore since we cannot see, hear, or touch them. While they are hard at work, many of us do not give much thought to the hidden minutiae of their constantly changing proprietary formulas. Their lines of complex and opaque code make lightning-fast decisions for and about us in both helpful and unhelpful ways.

Algorithms are not inherently good or bad. Rather, their effects depend on what they are programmed to do, who's doing the programming and to what end, how the algorithms operate in practice, how users interact with them, and what is done with the huge amount of personal data they feed on.¹⁸ On the plus side, these mysterious black boxes can answer in seconds a question that formerly required hours in a library (though the answer may not necessarily be entirely accurate). Social media platforms like Facebook, Twitter, and Instagram

let us share photos, personal news, and links with strangers across the globe whose interests align with ours. We can organize disaster relief or a grassroots social movement from far away. We can teach machines to pinpoint the location of brain tumors or help reduce traffic congestion.¹⁹

But algorithms also have influence we may not anticipate, since their use increasingly has political and societal dimensions.²⁰ Using incomplete datasets to predict odds of success, algorithms may determine who does and does not get into college based on their zip code rather than their academic efforts.²¹ Algorithms may be programmed to decide who is invited to interview and, ultimately, who gets a job offer.²² They might recommend which loan applicants are a good credit risk.²³ These invisible lines of code may even establish the length of a criminal sentence.²⁴

13 Cathy O'Neil (2016), *Weapons of math destruction: How big data increases inequality and threatens democracy*, Crown; Safiya Umoja Noble (2018), *Algorithms of oppression: How search engines reinforce racism*, New York University Press; Virginia Eubanks (2018), *Automating inequality: How high-tech tools profile, police, and punish the poor*, St. Martin's Press; Sofia C. Olhede and Patrick J. Wolfe (2019), "The growing ubiquity of algorithms in society: Implication, impact and innovation," *Philosophical Transactions of the Royal Society* 376(128), DOI: <https://doi.org/10.1098/rsta.2017.0364>

14 Zeynep Tufekci (2017), *Twitter and tear gas: The power and fragility of networked protest*, Yale University Press; Siva Vaidhyanathan (2018), *Antisocial media: How Facebook disconnects us and undermines democracy*, Oxford University Press; Yochai Benkler, Rob Faris, & Hal Roberts (2018), *Network propaganda: Manipulation, disinformation, and radicalization in American politics*, Oxford University Press; Jonathan Zittrain (23 July 2019), "The hidden costs of automated thinking," *The New Yorker*, <https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking>

15 John P. Wihbey (2019), *The social fact: News and knowledge in a networked world*, MIT Press, especially chapter 5, "Bias in network architecture and platforms"; Jihii Jolly (20 May 2014), "How algorithms decide the news you see," *Columbia Journalism Review*, https://archives.cjr.org/news_literacy/algorithms_filter_bubble.php; Laura Hazard Owen (15 March 2019), "One year in, Facebook's big algorithm change has spurred an angry, Fox News-dominated — and very engaged! — News Feed," *NeimanLab*, <https://www.neimanlab.org/2019/03/one-year-in-facebooks-big-algorithm-change-has-spurred-an-angry-fox-news-dominated-and-very-engaged-news-feed/>

16 Lee McIntyre (2018), *Post-truth*. MIT Press; Jonathan Zittrain (2019), op. cit.

17 For a list of keywords and definitions used in this discussion and throughout the report, see p. 49.

18 Hannah Fry (2018), *Hello world: Being human in the age of algorithms*, W. W. Norton & Company.

19 "Computer program beats doctors at distinguishing brain tumors from radiation changes" (16 September 2016), *Neuroscience News*, <https://neurosciencenews.com/ai-brain-cancer-neurology-5058/>; Francesca Baker (12 December 2018), "The technology that could end traffic jams," *BBC Future*, <http://www.bbc.com/future/story/20181212-can-artificial-intelligence-end-traffic-jams>

20 Lawrence Lessig (1999), *Code and other laws of cyberspace*, Basic Books; Mark MacCarthy (19 March 2019), "The ethical character of algorithms—and what it means for fairness, the character of decision-making, and the future of news," *The Ethical Machine*, Shorenstein Center, Harvard University, <https://ai.shorensteincenter.org/ideas/2019/1/14/the-ethical-character-of-algorithms-and-what-it-means-for-fairness-the-character-of-decision-making-and-the-future-of-news-yak6m>; Langdon Winner (Winter 1980), "Do artifacts have politics?" *Daedalus* 109(1), https://www.researchgate.net/publication/213799991_Do_Artifacts_Have_Politics

21 Bridget McCrea (16 November 2015), "Science develops an algorithm for college selection—but does it work?" *Ecampus News*, <https://www.ecampusnews.com/2015/11/16/algorithm-college-selection-329/>

22 Gideon Mann and Cathy O'Neil (9 December 2016), "Hiring algorithms are not neutral," *Harvard Business Review*, <https://hbr.org/2016/12/hiring-algorithms-are-not-neutral>; Drew Harwell (6 November 2019), "HireVue's AI face-scanning algorithm increasingly decides whether you get the job," *Washington Post*, <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/>

23 Cathy O'Neil, (2016), op. cit.

24 Julia Dressel and Hany Farid (2018), "The accuracy, fairness, and limits of predicting recidivism," *Science Advances* 4(1), DOI: <https://doi.org/10.1126/sciadv.aao5580>

In our daily lives, algorithms are often used to filter the news we see about the world,²⁵ potentially swaying decisions about what we buy and how we vote.²⁶ They may determine the results students get from searches in their college or university library.²⁷ At worst, data swept up by these algorithms can be used by state actors, criminals, or trolls bent on disruption or sabotage.²⁸

The big picture

The world of information has been transformed in unexpected ways in the past decade. These changes can be explained, in part, by the impact of algorithms. Figure 1 summarizes some of the factors driving these changes.

The impact of several convergent technologies and social trends helps explain how and why the world of information has changed so dramatically. A list of some of the most significant of these technological trends shows how these changes affect society.

1. Data collection is happening invisibly and constantly. We carry computers in our pockets that gather and share information about our daily lives, including where we go, who we associate with, what news catches

our attention, and what questions we ask. These data streams may be combined with information from data brokers²⁹ and harvested from our cars and household gadgets, like baby monitors, internet-connected thermostats, refrigerators, vacuums, and voice-activated assistants such as Alexa, Siri, and Google Home.³⁰

2. Advances in data science allow technologists and systems to collect and process data in real time, rapidly and on a vast scale (a development often called “big data”).³¹ This computational ability to quickly correlate enormous amounts of fine-grained, exhaustive data collected from numerous sources has opened up opportunities for companies and researchers — but also many Pandora’s boxes.
3. Automated decision-making systems are being applied to social institutions³² and processes³³ that, for better or worse, determine all kinds of things: who gets a job, a mortgage, or a loan, access to social services, admission to school or educational services.
4. Machine learning (ML) and artificial intelligence (AI), increasingly used in software products that make very significant decisions, often rely on biased or incomplete data sets. AI systems are “trained”³⁴ using

25 David R. Brake (2017), “The invisible hand of the unaccountable algorithm: How Google, Facebook and other tech companies are changing journalism,” *Digital Technology and Journalism*, Jingrong Tong and Shih-Hung Lo (Eds.), Palgrave Macmillan, 25-46.

26 Robert Epstein and Ronald E. Robertson (2015), “The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections,” *Proceedings of the National Academy of Sciences* 112(33), E4512-E4521, DOI: <https://doi.org/10.1073/pnas.1419828112>

27 Matthew Reidsma (11 March 2016), “Algorithmic bias in library discovery systems,” *Reidsma Working Notes*, <https://matthew.reidsrow.com/articles/173>

28 Bruce Schneier (2018), *Click here to kill everybody: Security and survival in a hyper-connected world*, Norton.

29 Molly Wood (7 March 2019), “Data brokers are in the business of selling your online information.” *Marketplace*, <https://www.marketplace.org/2019/03/07/when-it-comes-to-intrusive-data-collection-facebook-might-be-the-least-of-our-problems-data-brokers/>

30 Bruce Schneier (2015), *Data and Goliath: The hidden battles to collect your data and control your world*, Norton.

31 For a thorough introduction to the relationship of big data to knowledge, see Rob Kitchin (2014), “Big data, new epistemologies and paradigm shifts,” *Big Data & Society* 1(1), 1-12, DOI: <https://doi.org/10.1177/2053951714528481>

32 See for example Virginia Eubanks 2018, op. cit.

33 An accessible introduction to how algorithms influence society can be found in Cathy O’Neil 2017, op. cit.

34 Karen Hao (4 February 2019), “This is how AI bias really happens - and why it’s so hard to fix,” *MIT Technology Review*, <https://www.technologyreview.com/s/612876/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>

LIFE IN THE AGE OF Algorithms

DATA
are everywhere

collected 24/7

and processed in real-time & at vast scale

likes cats >>> more cats >>> cat food 50% off!

WSJ NYT FOX BUZZFEED DAILY NEWS BREITBART

Google f

Climate change is a flat-out hoax!

Will we survive climate change??

Platforms seek patterns in data to try to predict and influence behavior

and personalize news and advertising, based on guesses about individual preferences

creating the potential for us to experience different realities

job? yes.

mortgage? no.

prison sentence? 10 years.

INCOMPLETE/BIASED DATA + MACHINE LEARNING = THAT AMPLIFIES BIAS

PRECISION MAKING SYSTEM

YES!

NO!

Our lives are increasingly influenced by decision-making systems that build on correlations

using "artificial intelligence" software that relies on incomplete data to make generalizations

and amplifies existing biases at large scale

MERITOCRACY!

LIMITED PERSPECTIVE!

TECHNO UTOPIANISM!

SCALE!

that rely on mining users' data and manipulating their behavior

Information systems we depend on are shaped by a tech culture marked by narrow perspectives and overconfidence

in ways that undermine our trust in news, politics, and each other.

jessica yurkofsky (2019)

Figure 1: Life in the age of algorithms: A conceptualization

- existing, often human-edited³⁵ datasets, which means they can learn and amplify bias. This has implications such as teaching autonomous cars to avoid pedestrians³⁶ or recommending a prison sentence based on data from a criminal justice system that has a history of racial discrimination.³⁷
5. The disaggregation of published information and its redistribution through search and social media platforms makes evaluation of what used to be distinct sources, like articles published in an academic journal or stories in a local newspaper, all the more difficult. This disaggregation leads to an individualized presentation of information that sorts results based on inferences drawn from personal data trails. We do not all see the same information when we search and with original context missing, it is not obvious where it came from.
 6. There has been a rise of the “attention economy” or “surveillance capitalism”: profitable industries gather “data exhaust” from our interaction with computers to personalize results, predict and drive behavior, target advertising, political persuasion, and social behavior at a large scale.
 7. These industries appear to have difficulty anticipating or responding to unintended consequences. This may be because companies are influenced by Silicon Valley cultural values³⁸ that, among other things, consist of a belief in meritocracy, indifference to or ignorance of perspectives different from those of affluent White males,³⁹ a global reach coupled with a lack of cultural competence,⁴⁰ and magical thinking about the preeminent goodness of individualism and free speech.⁴¹
 8. Decades of media consolidation, deregulation, and economic trends combined with the rise of social media platforms that are designed for persuasion but have no ethical duty of care, have contributed to engineered distrust of established knowledge traditions such as journalism and scholarship, and the global destabilization of political and social institutions.

The technical infrastructure that influences how we acquire information and shapes our knowledge and beliefs has changed dramatically in ways that are largely invisible to the public — by design. We are facing a lack of public knowledge⁴² about who holds power over information systems and how that power is wielded, a gap in understanding that educators need to begin to address.⁴³ Given this sea change, questions about what it means

35 Mary L. Gray and Siddharth Suri (2019), *Ghost work: How to stop Silicon Valley from building a new global underclass*, Houghton Mifflin Harcourt.

36 Sigal Samuel (6 March 2019), “A new study finds a potential risk with self-driving cars: Failure to detect dark-skinned pedestrians,” *Vox*, <https://www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin>

37 Stephanie Wykstra (12 July 2018), “Can racial bias ever be removed from criminal justice algorithms?” *Pacific Standard*, <https://psmag.com/social-justice/removing-racial-bias-from-the-algorithm>

38 Alice E. Marwick (2013). *Status Update: Celebrity, publicity, and branding in the social media age*, Yale University Press.

39 Marie Hicks (2017), *Programmed inequality: How Britain discarded women technologists and lost its edge in computing*, MIT Press; Ruha Benjamin (2019), *Race after technology: Abolitionist tools for the new Jim Code*, Wiley.

40 Alexandra Stevenson (6 November 2018), “Facebook admits it was used to incite violence in Myanmar,” *The New York Times*, <https://nyti.ms/2yVAO5X>

41 Fred Turner (2010), *From counterculture to cyberculture: Stewart Brand, the Whole Earth Network, and the rise of digital utopianism*, University of Chicago Press; Timothy Garton Ash (2016), *Free Speech: Ten Principles for a Connected World*, Yale University Press.

42 The implications of this unequal “division of learning” are explored in Chapter 6 of Shoshana Zuboff (2019), *The age of surveillance capitalism: The fight for a human future at the new frontier of power*, PublicAffairs.

43 Annemaree Lloyd (26 September 2019), “Chasing Frankenstein’s monster: Information literacy in the black box society,” *Journal of Documentation* 75(6), 1475-1485. DOI: <https://doi.org/10.1108/JD-02-2019-0035>



What *exactly* is information literacy?

The term “information literacy” is sometimes mistakenly conflated with “library instruction,” but its meaning is really much broader. Information literacy is a collective effort of librarians, media specialists, technologists, and educators across the educational spectrum. It incorporates elements of media literacy, digital literacy, news literacy, and critical thinking.

Taken together, information literacy is an integrated set of skills, knowledge, practices, and dispositions that prepares students to discover, interpret, and create information ethically while gaining a critical understanding of how information systems interact to produce and circulate news, information, and knowledge.* ■

* The Association of College and Research Libraries (11 January 2016) has developed a “Framework for information literacy for higher education” that offers a discussion of the phrase and outlines core concepts for college-level students, <http://www.ala.org/acrl/standards/ilframework>

for students to be information literate today, and whether they know how information works in the age of algorithms, are of paramount importance.

Algorithms and higher education

How prepared are students for navigating a world of technologies that are fundamentally changing how we encounter, evaluate, and create information? Though they have grown up with internet giants such as Google, YouTube, Instagram, and Facebook, are students aware of these systems’ scale, scope, and velocity as their algorithms attempt to predict and influence behavior?

What role do these platforms play in their learning as students seek information they need for school and find their way through a thicket of daily news? What does “algorithmic justice” mean to a new generation of students affected by these systems but perhaps unaware that they are at work — even in their daily interactions with campus learning management systems, such as Canvas, online textbooks and advising and retention software?⁴⁴

Information literacy and critical thinking, dual competencies promoted on college and university campuses for decades, may come closest to addressing these weighty questions. The way these essential skills are taught, however, tends to concentrate on helping students meet their immediate academic needs: how to read texts closely and critically and how to use the web and library resources to find supporting materials for their assignments.

⁴⁴ Canvas, the market leader in learning platforms for digitally delivering readings, assignments, tests, and interaction between students and their instructors, announced its sale in December 2019 to a private equity firm for \$2 billion. In March 2019 Turnitin, a plagiarism detection company used on many campuses, was sold to a media conglomerate for \$1.75 billion. Quantifying students’ lives to improve student retention is also a growing movement. See Marc Parry (18 July 2018), “Colleges awakening to the opportunities of big data,” *The New York Times*, <https://www.nytimes.com/2012/07/22/education/edlife/colleges-awakening-to-the-opportunities-of-data-mining.html>

Drawing on 10 years of PIL studies of college students' research experiences and habits, we next identify the gaps in students' learning process and describe the skills and knowledge they will need to navigate information in the age of algorithms.

A decade of PIL research

In college, students practice habits of inquiry as they are asked to find and evaluate information sources in the campus library and online. And yet, a critical review of previous PIL studies from 2009 to 2018 suggests that students' approaches to research — and the challenges they face — have not changed significantly, and neither have the kinds of research-based learning opportunities faculty provide.⁴⁵ PIL's 2010 "Assigning Inquiry"⁴⁶ study found that most assignment instructions emphasized what the finished product should look like, and allowed only a narrow range of source materials, mostly peer-reviewed articles. Likewise, scholars have argued that academic paper-writing places "technical proficiency over intellectual depth."⁴⁷

Unsurprisingly, we found in PIL's 2010 "Truth Be Told" study⁴⁸ that students found it perplexing to figure out the nature and scope of the intellectual work instructors required of them. When left to their own devices, many took a familiar path, relying on the same sequence of steps to find "safe" sources regardless of topic, such as using

Google and Wikipedia to get started on research assignments before tapping library databases for acceptable sources. The same applied in everyday life for solving information problems, such as keeping up with news, making buying decisions, or checking out health and wellness information.⁴⁹

If the purpose of college research assignments is to prepare students to think critically and inquire deeply as they encounter new ideas, then assignments like these that encourage students to engage with information in such a limited way may miss the mark. Moreover, the information practices students develop to manage college assignments, according to our research, do little to equip them for an information environment that increasingly relies on manipulating large data sets to select and shape what they see. This is particularly true when it comes to learning how to evaluate information.

Students learn strategies for evaluating academic information to satisfy assignment requirements, but these may not transfer effectively to personal information seeking. While faculty help students gain the intellectual capacity to understand complex arguments made in scholarly books and journals through training in close reading and interpretation, this may come at the expense of equally important lessons.⁵⁰ Two PIL studies illustrate this worrisome disconnect between the critical information practices learned in college and the information skills students need in their

45 Research writing instruction appears to have taught students this approach to inquiry since at least 1961. See James E. Ford and Dennis R. Perry (1982), "Research paper instruction in the undergraduate writing program," *College English* 44(8), 825-831, DOI: <https://doi.org/10.2307/377339>. It is striking that the percentage of assignments that conform to a generic "research paper" was over 80 percent in surveys conducted in 1961, 1982, and in PIL's 2010 study of research assignment handouts.

46 Alison J. Head and Michael B. Eisenberg (13 July 2010), *Assigning inquiry: How handouts for research assignments guide today's college students*, Project Information Literacy, https://www.projectinfolit.org/uploads/2/7/5/4/27541717/pil_handout_study_finalvJuly_2010.pdf

47 Karen Manarin, Miriam Carey, Melanie Rathburn, and Glen Ryland, *Critical reading in higher education: Academic goals and social engagement*. Indiana University Press, 2015.

48 Alison J. Head and Michael B. Eisenberg (2010), *Truth be told: How college students evaluate and use information in the digital age*. Project Information Literacy, https://www.projectinfolit.org/uploads/2/7/5/4/27541717/pil_fall2010_survey_fullreport1.pdf

49 Alison J. Head and Michael B. Eisenberg (4 April 2011), "How college students use the web to conduct everyday life research," *First Monday* 16(4), <https://firstmonday.org/article/view/3484/2857>

50 A study conducted by the Stanford History Education Group found historians who relied on close reading of primary texts fared poorly when evaluating digital disinformation. Sam Wineburg and Sarah McGrew (6 October 2017), "Lateral reading: Reading less and learning more when evaluating digital information," Stanford History Education Group Work Paper No. 2017-A1, DOI: <http://dx.doi.org/10.2139/ssrn.3048994>

daily lives and after graduation.⁵¹ In 2012, PIL interviewed 23 U.S. employers who reported their new hires were inclined to rely on search engines for quick and superficial answers, had trouble seeing patterns and connections, and were reluctant to take a deep dive into a variety of information sources (see Figure 2).

In PIL's 2016 lifelong learning study, about three-quarters of college graduates believed that school had prepared them well to search for and analyze information. But only 27% agreed that college had helped them develop the ability to formulate questions of their own. Intellectual work for college assignments largely draws from and mimics the style of academic publications, which is profoundly dissimilar from the kind of open-ended and varied tasks they face after graduation. As one graduate said, "The faculty, the textbook author, they can carve the question any way they want, whereas in the real world it's not black and white, there's a lot of gray area and unknowns."⁵² Both PIL studies point to areas where graduates' research habits may make them more vulnerable to getting incomplete information as the top search results may reflect the priorities of algorithmic filters rather than the best information available.

Missing the mark

Students in this study confirmed what a decade of PIL research suggests: Their college assignments discourage them from working with information outside the realm of traditional scholarship. Moreover, their courses do not address the

51 See pp. 12-14 of Alison J. Head (16 October 2012), *Learning curve: How college graduates solve information problems once they join the workplace*, Project Information Literacy, https://www.projectinfo.org/uploads/2/7/5/4/27541717/pil_fall2012_workplacestudy_fullreport_revised.pdf

52 Alison J. Head (5 January 2016), *Staying smart: How today's graduates continue to learn once they complete college*, Project Information Literacy, https://www.projectinfo.org/uploads/2/7/5/4/27541717/staying_smart_pil_1_5_2016b_fullreport.pdf

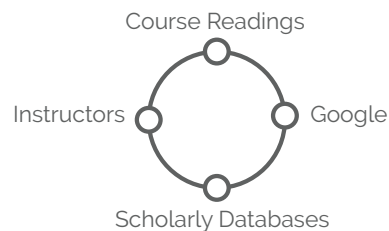
What Students Have Taught Us

A decade of PIL Research

The hardest part of research is getting started

Students also struggle with defining a topic, narrowing it down, and filtering what is irrelevant.

Students follow familiar paths through the same handful of trusted resources for research assignments.



Graduates are confident they can search and analyze but not that they can ask questions of their own



3 out of 4 students

College taught them to interpret and apply search results



1 out of 4 students

College helped far fewer formulate and ask questions of their own

Evaluating information is a collaborative process

61% turn to friends and family

49% turn to instructors

only 11% turn to librarians

Most students - 82% - think news is important for a democracy.

Yet, 36% said they don't trust any news no matter what source it comes from.

Figure 2: PIL findings on students' research habits, 2009-2018

significant social and ethical questions raised⁵³ about the workings of influential information systems on the public sphere. This is happening at a time when falsehoods proliferate and trust in truth-seeking institutions is being undermined. Even the very existence of truth itself has come into question. The classroom, however, seems strangely removed from this rising “post-truth” wave beating against the shores of our shared reality.

In PIL’s 2018 news engagement survey, eight in 10 students agreed news is “necessary in a democracy,” but journalism, most said, had fallen short of their idealistic standards of accuracy, independence, and fairness.⁵⁴ For them, staying current often meant navigating a minefield of misinformation, commercial interests, clickbait, “fast news” from social media, and political manipulation. Finding reliable information beyond the filter bubbles they knew constrained what they saw in searches and social media required work on their part. Strikingly, more than a third (36%) said “fake news” had made them distrust the credibility of *any news* at all.⁵⁵

The fact that a large proportion of students, who are bombarded by news on a daily basis, do not trust any of it indicates a large gap between the information literacy skills they practice for courses and their grasp of our current information environment. Particularly concerning is the fact that many students reported they were much more careful about selecting quality information for course assignments than they were for their personal consumption.⁵⁶ In many cases, students reported they were motivated to dig deeper when

a news story piqued their interest, but otherwise relied on well-informed friends or scanning a smattering of headlines in news digests to keep on top of events.

Why this matters now

PIL’s findings from the past 10 years come full circle when trying to understand how prepared today’s students will be to navigate an information landscape that has dramatically changed. These findings suggest that the shortcuts students adopt to manage academic research — and even the training they receive in carefully unpacking complex academic arguments — do not adequately prepare them for a world of abundant news and information that is deeply influenced by algorithms. These findings raise questions about whether the gap between what students learn in school and what they need to know is deepening at a time of an epistemological crisis.

Our research tells us that many of today’s traditional-aged students are different from those who came before.⁵⁷ Their ethnic and racial diversity, their professional destinies, and their experience growing up just as mobile devices and social media became ubiquitous, all set them apart.⁵⁸ To hear their stories, as we did in our 2019 focus groups for this study, many students think their experiences with technology distinguish them from other age cohorts. For instance, many were assiduously warned by concerned parents and teachers that the internet could be a “bad and dangerous place” where cyber bullying, human

53 See for example, Mark Bergen (2 April 2019), “YouTube executives ignored warnings, letting toxic videos run rampant.” *Bloomberg*, <https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant>

54 Alison J. Head, John Wihbey, P. Takis Metaxas, Margy MacMillan, and Dan Cohen (16 October 2018), *How students engage with news: Five takeaways for educators, journalists, and librarians*, Project Information Literacy, <https://www.projectinfolit.org/uploads/2/7/5/4/27541717/newsreport.pdf>, 14.

55 Op. cit. Head, et al. 2018, *How students engage with news: Five takeaways for educators, journalists, and librarians*, 15.

56 Op. cit. Head, et al. 2019, “Across the great divide.”

57 Rachel Premack (12 July 2018), “Millennials love their brands, Gen Zs are terrified of college debt, and 6 other ways Gen Zs and millennials are totally different,” *Business Insider*, <https://www.businessinsider.com/gen-zs-habits-different-from-millennials-2018-6>

58 danah boyd (9 March 2019), “You think you want media literacy...do you?” *Points*, Data & Society, <https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2>

trafficking, and predators congregated to spread their destructive messages on an epic scale.

Searching on school-issued laptops became a supplement to (or in some cases, a replacement for) using school libraries.⁵⁹ And, while teachers were giving these high schoolers ineffective lessons, such as “domains ending in .org are more trustworthy than those ending in .com,” and relying on standard checklists to evaluate websites,⁶⁰ students were teaching each other how to circumvent the filters schools used to keep them from landing on the “wrong” websites. In the process, they were learning how to insulate themselves from the surveillance and control of parents and teachers.

Today, this cohort is coming of age at a time when electoral politics are playing out through social media memes, and Twitter feuds become tomorrow’s headlines. Social media platforms are used to stage massive student walkouts against gun violence and a cross-generational “March for Our Lives” around the world. Turning Point USA,

a conservative group, calls out faculty advancing “leftist propaganda in the classroom” with “The Professor Watchlist” website.⁶¹ A 17-year-old Swedish climate activist⁶² rallies people across the globe through social media and exchanges online barbs with the U.S. president. It is clear that today’s young people will help determine what is possible as a collective society.

Many college students already see themselves as active participants in news and information flows, not passive recipients of uncontested knowledge. Our work as educators and librarians is to help students navigate information, not just for college courses but beyond — in the workplace, in their personal lives, as lifelong learners, and as news consumers, creators, and voters. The ways information is shaped and shared today has changed a great deal since we began our national research studies at PIL only a decade ago. Taken together, this means we need to change how and what we teach. But first, we must understand what students may already know about algorithms.

59 Hallie Golden (4 September 2019), “The decline and evolution of the school librarian,” *CityLab*, <https://www.citylab.com/life/2019/09/school-libraries-books-literacy-librarian-jobs-education/597316/>

60 Students in this study mentioned lessons learned in high school, generally in dismissive terms. For a study that faults the chronic problem of ineffective web training in schools, see Joel Breakstone, Mark Smith, Sam Wineburg, Amie Rapaport, Jill Carle, Marshall Garland, and Anna Saavedra (14 November 2019), “Students’ civic online reasoning: A national portrait,” Stanford History Education Group. <https://stacks.stanford.edu/file/gf151tb4868/Civic%20Online%20Reasoning%20National%20Portrait.pdf>

61 Christopher Mele (28 November 2016), “Professor Watchlist is seen as a threat to academic freedom,” *The New York Times*, <https://www.nytimes.com/2016/11/28/us/professor-watchlist-is-seen-as-threat-to-academic-freedom.html>

62 Somini Sengupta (18 February 2019), “Becoming Greta: ‘Invisible girl’ to global climate activist, with bumps along the way,” *The New York Times*, <https://www.nytimes.com/2019/02/18/climate/greta-thunberg.html>

Part Two:

Student Perceptions and Concerns

Everyone who has accessed the internet has experienced the personalizing actions of algorithms, whether they realize it or not. These invisible lines of code can track our interactions, trying to game our consumer habits and political leanings to determine what ads, news stories and information we see. Companies, using algorithms like this, work nonstop to amass behavioral data on every aspect of our lives that they can combine, use, and sell.⁶³ As tracking practices have become more common and advanced, it has become urgent to understand how these computer programs work and have widespread impact. How do students understand the hidden filters that influence what they see and learn, and shape what they think and who they are?

In this section of the report, we present empirical data that explores college students' awareness and perceptions of algorithm-driven platforms, and what actions they take, if any, to safeguard their privacy.⁶⁴ Qualitative data were collected from 16 student focus groups with 103 college sophomores, juniors, and seniors, and from telephone interviews with 37 faculty members at eight U.S. colleges and universities.

The findings are presented as four takeaways that detail and discuss how students experience algorithms — their individual perceptions, experiences, concerns, and opinions as well as consensus about these things among participants. These takeaways serve as the basis for actionable recommendations later in this report to guide stakeholders considering possibilities for preparing students for the future.

» ***Takeaway 1: Students have an ambivalent bond with algorithm-driven platforms.***

Almost all of the students in our focus groups were aware that platforms pushed content onto their screens. While most said they had no idea how algorithms actually worked, they had definite opinions about the effects personalization had on their online lives. As one self-described novice explained, “I’m really not a guru, but search engine algorithms take what you click on and they make this magical potion that always caters to sites you constantly use for news like CNN instead of BBC.”

Students were quick to give examples of targeted advertising, where the inner logic of algorithmic persuasion was most visible as the same ads chased them across platforms and devices. They often swapped stories about how some personalization attempts had fallen flat. One student said that a search for political internships in D.C. yielded an internship in Korea. Another said she kept getting ads for flights to a place she had just visited.

While discussions often began with reactions to advertising behavior, students also talked about algorithms used for personalizing other types of content. According to students, the use of algorithms by social media platforms was pernicious. As one said, sites like Facebook “can serve information and news, almost like propaganda, they can serve up whatever information they want to, and potentially change your worldview — it can be a double-edged sword.”

63 Bennett Cyphers (2 December 2019), “Beyond the one-way mirror: A deep dive into the technology of corporate surveillance,” Electronic Frontier Foundation, <https://www.eff.org/wp/behind-the-one-way-mirror>

64 For another view of how youth think about digital privacy, see Helen Creswick, Liz Dowthwaite, Ansgar Koene, Elvira Perez Vallejos, Virginia Portillo, Monica Cano and Christopher Woodard (2019), “‘...they don’t really listen to people’: Young people’s concerns and recommendations for improving online experiences.” *Journal of Information, Communication and Ethics in Society* 17(2), 167-182, DOI: <https://doi.org/10.1108/IICES-11-2018-0090>

Getting news online through the internet giants revealed a growing tension between receiving the services they wanted and the psychic cost. One student, resigned to the tradeoff, admitted, “I’m giving up control of my personal data so I can get news I agree with, see my interests reflected in my feed, and shop more efficiently.”

Resignation and indignation

An important takeaway from our focus groups was the profound ambivalence — the tangle of resignation and indignation — that almost all students expressed about algorithm-driven platforms that collect data about their personal lives. While many students objected to certain advertising practices platforms used, they were nonetheless resigned to using sites like Google and YouTube. In their words, algorithms were “part of the deal” if they wanted to use “free apps.” Fewer, however, seemed to realize that their motivations to connect and get content could be exploited in ways that extended beyond targeted ads. While some said they knew Google and Facebook were “advertising companies” and “their goal is to tailor advertisements to you,” most found these sites too useful to abandon. Still for others, algorithms were necessary and welcome, since personalization helped filter the vast number of irrelevant Web results they might otherwise get from a Google search. As one student summed it up about his cohort, “We would rather have this convenience than protect our privacy.”

Comments like these suggest students have an awareness about the benefits of algorithms for sorting information to surface relevant results. At the same time, they signal students’ awareness of the pitfalls of compartmentalizing people to influence their actions and exacerbate divides. But

exactly what a company might be doing with the data collected about them was often an unknown variable in students’ cost-benefit analysis. While some weighed whether the ability to use certain sites was worth sacrificing their data, many others claimed it was already too late. As one student rationalized, “Your information is 100% out there somewhere, and it’s definitely concerning, but I want an Amazon account, so I can purchase things, and I want a Facebook account, too.” There was widespread consensus by students that these sites helped them stay in touch with friends and family. As one student admitted about his ambivalence:

I just feel desensitized to it. If I really think about it, I’m like, yeah, that’s messed up, and I don’t want them listening to me, but then I still use those apps and am letting it happen. So obviously it doesn’t matter to me that much. I mean what are you gonna do?

Contradictory feelings such as these were most evident when students were asked about their “tipping points.”⁶⁵ When had an algorithm-driven platform had gone too far? The potential for conversations to be picked up by devices and trigger related advertising was a common tipping point for them (see sidebar, “When algorithms ‘get creepy’”). Some students said they turned off their device’s microphone to retaliate against the practice; many others said that they refused to have Alexa in their residences. In the words of one student, “the phone is already doing enough listening.”⁶⁶ Another added, “It’s interesting that people who are most close to working technologies like Alexa don’t use them.”

Others questioned how wedded they were to online behemoths after learning Cambridge Analytica used algorithms to develop a targeted campaign to

⁶⁵ In Malcolm Gladwell’s bestselling book, *The tipping point: How little things can make a big difference* (2000), the author describes the tipping point as “the moment of critical mass, the threshold, the boiling point,” p. 304.

⁶⁶ While a concern about phones “listening” is widespread, it’s most likely not happening routinely. See Bree Fowler (10 July 2019), “Is your smartphone secretly listening to you?” *Consumer Reports*, <https://www.consumerreports.org/smartphones/is-your-smartphone-secretly-listening-to-you/>



When algorithms “get creepy”

The rich discussions in our focus groups showed that students across the country had some common concerns about algorithms.*

In an analysis of logs and transcripts, the PIL Team used eight individual themes based on whether one or more students in one of the 16 groups ($N=103$) had raised the concern (see Table 1). If students mentioned a concern, for instance, about “the next generation” more than once in a session, we only counted it once.

More than anything, we heard concerns from students about the “creepiness” of algorithms that violated their privacy. This happened when platforms ‘overheard’ their discussions or shared data with each other to pitch them products.

I was having a conversation with my friend and joking about somebody being pregnant, and then ads started popping up on my searches for pregnancy tests and supplements. I was laughing because Google got it wrong, but it’s still creepy.

While students in many groups worried about how the next generation would fare, they, themselves, were often unfamiliar with how the use of algorithms was developing and expanding. Automated decision-making that could directly affect their lives was particularly disturbing. They also remarked on the disappearance of a shared reality that results from personalized news and information. In many discussions these societal and personal concerns intersected. ■

* PIL researchers Alaina Bull and Jessica Yurkofsky did the coding for this content analysis from 10 October – 4 December 2019.

Table 1: What worries students about computer algorithms?

Concerns about Algorithms	Count	Percent
1. (P) Platforms “listening” across devices or platforms.	14	88%
2. (S) Algorithms & automated decision-making reinforcing inequalities.	12	75%
3. (P) Platforms shaping individual content & ads they see.	12	75%
4. (S) Online users not seeing the same reality.	11	69%
5. (S) The next generation.	10	63%
6. (P) Platforms selling personal data to third parties.	8	50%
7. (P) Permanence of data being collected about them.	7	44%
8. (S) Older generations using these technologies & adapting to changes.	5	31%

*Count is based on concerns discussed per 16 student focus groups.
(P = Personal concerns, S = Societal concerns)*

millions of Facebook profiles to sway users' voting habits. Still others referred to "deep fakes"⁶⁷ and viral hoaxes, such as a video clip of House Speaker Nancy Pelosi deliberately slowed to make her appear inebriated⁶⁸ that spread through Facebook, YouTube, and Twitter in May 2019. Some said Facebook's refusal to take the video down, along with widespread sharing of the video by pro-Trump supporters, had crossed the line for them.

At the same time, many students shared concerns about echo chambers that deepened social inequalities. A woman of color related algorithmic injustice to the biases that led to over-policing Black communities while failing to prevent mass shootings by Whites from affluent communities. Another student described predictive algorithms as "just a fancy technology-enabled form of stereotyping and discrimination, which is inherently problematic, but it's easier for us to overlook because it's happening online and we're not seeing it." Still another student related how he had thought about majoring in the field of computer science, but decided against it:

I don't like the direction that technology is going. A lot of it can be used for evil, and even though it's really smart, and it's like really well implemented and effective for the people who it's serving, it's not serving the general population. And that freaks me out.

Together, these comments suggest that many students, though not all, have a broader understanding of the impact of algorithms on their lives beyond the strictly personal effects of advertising that many were eager to discuss. They were torn by the appeal of using "free sites" while knowing they were being tracked for undisclosed purposes. And almost all of them were still trying to figure out some way to circumvent online surveillance and shield their vulnerability,

regardless of how effective their methods may have been. As one student described a sense of resignation combined with indignation: "It's a horrible totalitarian hellscape, but it's kind of the best we can reasonably expect."

» ***Takeaway 2: Students use defensive practices to protect their privacy.***

Often out of frustration, some students had taken to "gaming the algos," a combination of practical strategies to protect their privacy that we came to call *defensive practices*. These practices included using free apps and browser extensions to counter tracking along with a few retaliatory tactics they had invented or learned from friends to intentionally "confuse algorithms." While some students were adamant that their strategies were effective, many more were unsure about the net effects their actions were having.

Most commonly, students were running ad blockers or had regularly cleared their browsers of cookies, two old-school tactics for protecting their privacy. A few said they used DuckDuckGo as a search tool or Firefox as a browser to protect their privacy, countering consolidation of personal information by Google, and trying to avoid being put into the filter bubble of personalized search results. Others said they ran Virtual Private Networks (VPNs) to shield their browsing activity. Still others created multiple accounts on platforms like YouTube, Google, and Instagram so they could avoid having all of their internet activities tied to a single identity. As one student explained, "I got irritated because it sees who I am, and then it funnels me into getting certain content."

For some students being reduced to a profile based on the profligate collection of data created a problem of "context collapse" — they lost control

67 Oscar Schwartz (12 November 2018), "You thought fake news was bad? Deep fakes are where truth goes to die," *The Guardian*, <https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth>

68 Charlie Warzel (26 March 2019), "The fake Nancy Pelosi video hijacked our attention. Just as intended," *The New York Times*, <https://www.nytimes.com/2019/05/26/opinion/nancy-pelosi-video-facebook.html>

of the ability to craft their public identity when different facets of their personality became jumbled in a single account. They solved this problem by deliberately curating multiple “selves,” or different accounts that could reflect their varied interests. This defensive practice was intriguing because it gave them control of their self-representation as they interacted with the digital communities that formed around each of those interests, engaging in a kind of digital code switching.

International travel emerged as a surprising catalyst for learning about the role of algorithms in personalizing content. While traveling abroad, students said they were forced to learn new tools and strategies like using VPNs to get around firewalls. In some cases, they described seeing firsthand how international events could be misrepresented in the mainstream Western news. It also brought home the reality that geographic location was a major factor in personalizing search results and other content.

But not all students in our sessions said they were taking action to protect their identity from prying algorithms. These were often the optimistic students who said algorithms did more good than harm. A trend that emerged was that those who discussed using one defensive practice in the focus group ended up talking about using several. The students who were most aggressive in their strategies were often STEM majors, or lived with roommates who were STEM majors and were a source of useful tactics. While some students were clearly more knowledgeable than others about how to counter online tracking attempts, there was consensus throughout the focus groups that using sites like Google, Facebook,

YouTube, or Instagram had made them much more vulnerable to tracking.

No matter their major, it was clear that most students wanted to learn how to fight back against online surveillance. In fact, when the topic came up in the focus groups, it was not unusual to see students jot down a few notes about apps, like AdBlock, or server subscriptions, like NordVPN, that fellow focus participants said they were using. We found this tendency to learn from peers had likely developed earlier during adolescence as they evaded parental oversight through digitally mediated peer relationships,⁶⁹ and as they shared hints about navigating around the school’s imposed barriers. As one student explained, “Everyone just kind of shared this information around the entire school, like, ‘Oh, get the VPN and it’ll hack everything and you can access your Instagram during the middle of the day.’”

Comments like this one suggest that many students, depending on their socioeconomic status and high school resources, may enter college already knowing far more about navigating the internet than many might suspect. And yet, many students in this study were largely unaware that systems they used in their courses, like Canvas, the popular learning management system (LMS), had the potential to gather, aggregate and sell personal information. Once it was discussed in our sessions they were indignant to learn it might be happening. Though university administrations often claim that such surveillance is valuable for student retention and assessment, some educators are concerned about “learning analytics” programs that fail to offer students opportunities to provide

69 danah boyd (2014), *It’s complicated: The social lives of networked teens*, Yale University Press.

informed consent or opt out.⁷⁰

Breaking the news bubble

Targeted ads were clearly an annoyance, but “filter bubbles”⁷¹ were even worse. Most students knew that algorithms showed them only part of the picture, especially on social media platforms like Instagram and YouTube. In turn, these personalization effects, these “silos,” or “echo chambers,” had trapped them in a narrow space of confirming and reinforcing beliefs. One student expressed frustration by saying, “Just because I watched a video on whatever subject doesn’t mean that I don’t want to see the opposing side, I want to be educated, I don’t want to be in my box with one opinion.” Another described concern about the silo effect of social media platforms:

I often worry about getting everything because I usually ‘like’ pages that I agree with or ‘follow’ pages that I agree with, but that actually worries me because I feel like it will put me in this bubble where I don’t have any exposure to different opinions, so I made a conscious effort to not unfollow pages I dislike.

Many students, like this one, tried to evade news traps, so they could see “other sides of news” and escape the perils of personalization. Often

they used what academics call *lateral reading*,⁷² a strategy for seeking out sources that would present different approaches to the same topic. For example, students might compare how the same news story was covered by one source, such as *The New York Times*, and by more conservative sources, like *Breitbart News* or Fox News. One student explained reading across content producers to get the complete story:

I don’t trust one source — I purposely follow the other side, I guess it’s weird, but I want to see how they’re thinking, too, because it gives me some insight on how they’re forming this article or opinion.

Though many agreed that lateral reading was the best strategy to ensure a balanced view and check for accuracy, others were clearly exasperated with the amount of work that they had to put into this process:

I see something on Facebook and I try to get the true information. I’m like, ‘Okay, how do I know which websites have not just posted a bunch of bull, how many websites do I need to scroll through to find what I’m looking for?’

A few students said they were taught in college to trace scientific research reported in the news to its source. Still others, critical of the incompleteness and inaccuracy of breaking news coverage, used local news sites to get “the real story,” with some

70 Some students on campuses outside our sample, however, are becoming aware of data-gathering by institutions and are increasingly advocating resistance. See Zak Vescera (27 March 2019), “Canvas is tracking your data. What is UBC doing with it?” *The Ubyyssey*, <https://www.ubyssey.ca/features/double-edged-sword/>; Tom Nash (November 2019), “Freedom of Information Request: Ram attend opt out,” Muckrock, <https://www.muckrock.com/foi/virginia-128/ram-attend-optout-83757/>; Sharon Slade and Paul Prinsloo (2014), “Student perspectives on the use of their data: Between intrusion, surveillance and care,” *European Distance and E-Learning Network*, 18(1), 291–300, <https://www.eurodl.org/index.php?p=special&sp=articles&inum=6&abstract=672&article=679>. Also see Kyle M.L. Jones (2 July 2019), “Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy,” *International Journal of Educational Technology in Higher Education* 16(24) DOI: <https://doi.org/10.1186/s41239-019-0155-0>; Kyle M.L. Jones and Dorothea Salo (April 2018), “Learning analytics and the academic library: Professional ethics commitments at a crossroads,” *College & Research Libraries* 79(3), 304–323, DOI: <https://doi.org/10.5860/crl.79.3.304>; Deborah West, Ann Luzeckyj, Danny Toohey, Jessica Vanderlelie, and Bill Searle (1 January 2020), “Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics,” *Australasian Journal of Educational Technology*, DOI: <https://doi.org/10.14742/ajet.4653>

71 Eli Pariser popularized this concept in 2011 in *The filter bubble: How the new personalized web is changing what we read and how we think* (Penguin). However, critics have since called the significance of filter bubbles into question. See for example Mario Haim, Andreas Graefe, and Hans-Bernd Brosius (2017), “Burst of the filter bubble? Effects of personalization on the diversity of Google News,” *Digital Journalism* 6(3) DOI: <https://doi.org/10.1080/21670811.2017.1338145> and Frederik J. Zuiderveen, Damien Trilling, Judith Möller, Balázs Bodó, Claes H. de Vreese, and Natali Helberger (31 March 2016), “Should we worry about filter bubbles?” *Internet Policy Review* 5(1), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2758126

72 For a discussion of lateral reading, see op. cit. Sam Wineburg and Sarah McGrew 2017, 38–40.

waiting days — even weeks — to read a breaking news account so they could find out what had really happened.⁷³ As one student argued, it was the responsibility of news consumers to dig deeper and go beyond social media news feeds:

People shouldn't just rely on social media to get news, they should hold themselves to a higher standard, and if they see something, then they should look deeper and try to find out if it's true or false. So, I think that people should just make their own decisions based on their own research, instead of just looking at social media posts and just agreeing with it, without really thinking too much.

Regardless of how they got news and information, most students were especially concerned about how algorithms tailored information to the individual in ways that reinforce beliefs, biases, and prejudices. For some, this is how algorithms went beyond simply being code and preyed on human nature. Still other students acknowledged they were willing to be categorized so they could have their thoughts, opinions, and news preferences confirmed.⁷⁴

As one student pointed out, this is nothing new: “We surround ourselves with people we agree with and then we have conversations with them that reinforce our own ideas, so in a lot of ways, this is now just getting a third party to do all of that for us.”

» ***Takeaway 3: Trust is dead for many students, and skepticism lives.***

The way information is delivered today, with political propaganda and opinion mingled with traditional news sources, and with algorithms highlighting sources based on engagement

potential rather than quality, many students expressed concern about the trustworthiness of online content. Almost all of the students in our focus groups were critical of the current news environment. Some complained about the difficulty of knowing where to place trust. Another put it this way, “You’re getting people’s opinion on what happened versus what actually happened.”

At the same time, several students in each focus group were cynical almost to the point of believing their concerns and actions had little meaning, and that it was not possible to change things.⁷⁵ These students usually questioned the trustworthiness of news and information online but within a broader global context.

We’re kind of a nihilistic generation, we’re more existentialist, more so than the younger generation after us. We have a feeling kind of like, we’ve basically got a fascist leadership of the country, the climate is screwed, and I could go on for an hour about the millions of problems that we’re facing, so there’s this feeling like, whatever, we’ll just suck it up.

Given the choice

The theme of choice was mentioned repeatedly in our sessions. That seemingly simple word was filled with complexity and nuance when used to discuss perceptions of online content today. Some students said they had too many choices in navigating the crowded news landscape, where credible coverage is mixed in with a deluge of poorer-quality online content and misinformation. According to students, accepting algorithmic sorting was complicated by the number of news channels and sources now available, and this required

⁷³ For more about growing issues with the veracity of breaking news, see op. cit., Head, et al. 2019, “Across the great divide,” p. 28.

⁷⁴ Jeff Passe, Corey Drake, and Linda Mayger (July 2018), “Homophily, echo chambers, & selective exposure in social networks: What should civic educators do?” *The Journal of Social Studies Research* 42(3), DOI: <https://doi.org/10.1016/j.jssr.2017.08.001>

⁷⁵ This finding is consistent with PIL’s 2018 news study. Op. cit. Head, et al. 2018, *How students engage with news: Five takeaways for educators, journalists, and librarians.*

additional work to parse out:

People have an autonomy and a free will to participate in the distribution of information. And it's not that we're lacking credible information. It's that we're drowning in like a sea of all these different points out there, and people are willingly giving themselves up to participate in that sea.

Some students had strategies for navigating the plethora of choices by relying on crowdsourcing to keep up with important news. They used Reddit to point them to the most important headlines of the day. Others preferred to curate their news feed by selecting who to follow on Twitter. As one student described it, these social media sites functioned like a news editor. Still other students pushed back against algorithmic tailoring of news content, saying it amounted to “taking away personal choice unless you make a new account, like a blank slate, of your internet personality.”

One student highlighted broader social concerns, by seeing the potential of predictive algorithms to reduce choice; presenting the illusion of personalization while actually reinforcing a more homogeneous view of the world:

I'm more concerned about, like, the larger scale trend of predicting what we want, but then also predicting what we want in ways that push a lot of people towards the same cultural and political endpoint. I feel like that's talked about less than, like, individual privacy aspects.

As this student insightfully pointed out, there are social harms beyond the loss of personal privacy that have the potential to influence society at scale. An example is the ways extremists have promoted radical ideas to a wide audience (see sidebar, “The mainstreaming of extremism”). Altogether, issues around choice and agency appear closely related to trust and skepticism and

to students' suspicions about being manipulated by invisible forces.

These findings suggest that taking a skeptical approach to all information has become a reflex for many students, with many considering lateral reading as the default defense. Though one student said sticking to “big trustworthy sites” like CNN and *The New York Times* was the best way to navigate news, another student took a more cynical stance, asserting, “I don't consider any news source to be credible anymore.”

Schooled for skepticism

An important theme to emerge from our sessions was that no news source could be trusted at face value. This viewpoint did not appear to be a symptom of political partisanship so much as being a pervasive belief among students that they should rely on themselves to decide what to believe.⁷⁶ Students attributed this outlook to having come of age as the web has evolved from a collection of dubious websites to a dominant news portal and focal point for their social lives. Many felt they had been schooled to be critical of everything they encountered.

One student suggested skepticism in her age cohort accompanied the growth of the internet:

When we first started, we didn't have to filter through what was a credible source, and now you kind of got to filter through everything. We have different eyes as we're looking at everything, like literally everything on the internet. We're skeptical.

Still another student suggested they were disposed to doubt even the authority of their teachers: “We're all super cynical and untrusting of information to the point that we want to find

⁷⁶ In the 2018 PIL News Study, political affiliation was found to be positively correlated with distrust of news, as it is in the general population. However, in this qualitative study most, though not all, students discussed skepticism as a trait they shared without regard to their political beliefs. Op. cit. Head, et al. 2018, *How students engage with news: Five takeaways for educators, journalists, and librarians*.



The mainstreaming of extremism

Students in the focus groups were aware of how popular platforms used algorithms to shape what information they received while utilizing their engagement – their clicks and likes and shares – for profit.

What they were less aware of were the details of how companies like Facebook harness engagement by giving customers, including those promoting radical ideologies, the tools to target individuals using a menu of attributes including ethnicity, income, political orientation, hobbies. Investigative journalists have even found filters that enable clients to promote content to certain fringe factions, such as “Jew-haters” or fans of Joseph Goebbels.* These invisible fine-grained filters enable extremists to find and cultivate potential allies.

Students are not immune to such extremist appeals. During our interviews, a professor said she had been approached about a classmate posting messages about White supremacy with swastikas on his social media newsfeed. This situation was resolved, but the instructor was shaken by the incident: “We’re becoming a much more heterogeneous society in terms of worldviews and beliefs because access to that information is so easy to find to support your worldview.”

Technology is not the sole culprit in the amplification of fringe ideologies. There is a perfect storm brewing of the news industry, the attention economy, and coordinated actions of certain ideologues that have coalesced to drive extremist views into public prominence. White supremacists have become adept at harnessing the power of virality to find susceptible audiences and push their ideas and conspiracy theories into the mainstream. They have capitalized on long-term trends undermining trust in government and in those truth-seeking institutions we once turned to for authority: journalism, science, and the academy.†

Fixing the mechanisms that amplify distrust will not eliminate the underlying structural drivers of extremism. Students need to learn about the social and historical context of extremist beliefs as well as how extremism is mainstreamed technologically. Information literacy is more than knowing how to use technology and which buttons to push. It must also address how our emotional buttons are pushed, who is doing the pushing, and why. ■

* Sam Dean (21 February 2019), “Facebook decided which users are interested in Nazis - and let advertisers target them directly,” *Los Angeles Times*, <https://www.latimes.com/business/technology/la-fi-tn-facebook-nazi-metal-ads-20190221-story.html>; Julia Angwin, Madeleine Varner, and Ariana Tobin (14 September 2017), “Facebook enabled advertisers to reach ‘Jew haters,’” <https://www.propublica.org/article/facebook-enabled-advertisers-to-reach-jew-haters>

† Jessie Daniels, Mutale Nkonde, and Darakhshan Mir (May 2019), “Advancing racial literacy in tech: Why ethics, diversity in hiring, and implicit bias trainings aren’t enough,” *Data & Society*, https://datasociety.net/wp-content/uploads/2019/05/Racial_Literacy_Tech_Final_0522.pdf; Yochai Benkler (17 October 2019), “Cautionary notes on disinformation and the origins of distrust,” *MediaWell*, <https://mediawell.ssrc.org/expert-reflections/cautionary-notes-on-disinformation-benkler/>

it out ourselves, so if a teacher says, ‘There’s five rows, then we actually look, and, yep, there’s five rows.’” This skepticism, as one student described it, was a generational trait.

It’s different between students and professors, because they come from a pre-social media age and they’re used to being able to trust kind of different resources that they’ve always gone to. Whereas we grew up with untrustworthy sources and it’s drilled into us you need to do the research because it can’t be trusted.

As a whole, we found that the lack of trust in traditional authority figures meant trust was placed in Google as the arbiter of truth, sometimes to a ridiculous extent. One student who is also a parent described how he had tried to explain to his child that the bogeyman was not real, but his child had not believed him until a Google search confirmed it.

Some students said learning to approach all information critically was a valuable feature of their college education. And yet, they seemed to distinguish the reflexive skepticism they developed when sorting through websites for high school projects from the kind of critical thinking encouraged in college. This practice of discernment in college involved analysis of complex texts as well as applying social and historical context to current events, as they often did when teasing out the social implications of algorithms.

» ***Takeaway 4: Discussions about algorithms barely, if ever, make it into the classroom.***

Despite students and faculty expressing deep concern about the ways algorithms shape and influence how we learn about the world, it was surprising to learn this topic rarely came up in the college classroom. Instead, many students, like amateur sleuths, had discovered algorithms through keen observation, noticing how their content was personalized and different from what their friends were seeing.

Students’ suspicions were often confirmed in informal discussions with friends or relatives. As one student said, “A lot of times I learn about technology by copying what other people are doing, like my cousin was using a VPN, so I started using one too.”

When we asked instructors about the courses they teach, and if they had helped students think about how information is created and encountered on different online platforms, the responses were telling. Of the 37 faculty interviewed, only 10 came up with answers that addressed algorithmic platforms. Most strategies they mentioned were superficial, such as introducing students to DuckDuckGo as an alternative to Google when searching the Web.

One professor described a case study that used news reports of an event in the past to demonstrate how news is socially constructed. He encouraged students to look not at whether the claims made are true or false, but “what social and cultural work those truth claims are doing.” Still, though, the professor did not explicitly tie that to the current information landscape to explore how algorithms could filter meaning.

Instead, it was far more common for faculty to not even consider the changing information landscape. These instructors often expressed societal concerns about personalization, mourning “the loss of a common culture” and pointing to deeper epistemological implications, but it did not influence their teaching. As one instructor put it, “The genie is being let out of the bottle and it is a problem we’re stuck with now.”

The large majority of faculty members said they saw value in encouraging students to use peer-reviewed research and instilling critical thinking practices through close reading and textual analysis. In a few cases, faculty said they had scheduled a visit from a librarian in their courses to provide information literacy lessons.

One faculty member seemed surprised that it had never occurred to her:

I'm concerned societally that lots of us are getting sort of one perspective on events, and it isn't being presented, as you know, the most accurate perspective on events. And that we would all benefit from seeing a wider variety, or having more sort of similarity in the things that we see but I've never, that I can think of, specifically talked about the way that that search algorithms or algorithms that are providing news might, might have an influence.

Others thought algorithmic literacy should be part of a college education — so long as someone else took the lead.

It probably should be happening in every classroom space, or be sort of introductory things required of all students entering higher ed. Ideally, that would be really useful as an instructor who's working with juniors and seniors, I would like to not have to teach those skills.

And yet, when asked, this instructor, like most, said she did not discuss how algorithms influenced the information environment in her courses. An important finding to emerge about faculty is they thought a greater attention to generic critical thinking and rhetorical analysis skills early in the four years of college would prepare students to navigate the current environment. Indeed, students at one institution pointed to a required course on critical thinking as the place where such discussions belonged.

But many instructors seemed to assume critical thinking skills taught traditionally were sufficient. "With critical thinking," one instructor said, "you have to teach them how to think for themselves and how to pose certain questions to themselves about the universe and their place in it and using scientific inquiry, no matter what it is that you're tackling, no matter what your research project." Since critical thinking is part of every

discipline, one instructor argued, there was no pressing need for change.

Some students agreed. At one focus group, they said their institution valued critical approaches, and because it was such a strong institutional value, they were confident they could apply those skills widely (though they had previously expressed concern and a sense of helpless inevitability when it came to the workings of algorithms). Students elsewhere were more likely to draw no particular connection between what they were asked to do for school and the kinds of information practices they needed for everyday life. As one stated, "My instructors usually encourage us to use the databases provided by the school that can let you know that you're getting academic journals or peer reviewed journals that are solid information."

Essentially, students thought algorithmic systems were part of life, but the information students needed for school assignments had nothing to do with life beyond college. Some scoffed at the outdated advice about the internet provided in their courses: "They talk a lot in school about .org or .edu. But now with YouTube videos, they don't really have those things." Because most students believed that they knew more than their teachers about algorithmic technologies, they saw no value in addressing them in the college classroom. In one focus group, students agreed that their professors were too clueless about technology to cover it in courses, but they "forgave" their professors because they had other valuable knowledge to share. In another group a student was dismissive:

Usually, it's like a two-day thing about 'This is how you make sure your sources are credible.' Well, I heard that in high school, you know, and that information is just kind of outdated for the caliber that the internet is today. I mean it's just not the same as what it used to be.

Despite being alarmed about trends in technology, students found personal experience to be a better teacher than anything that might happen in a college classroom. Many expressed a preference for learning from peers. Yet, students were intrigued by the implications of algorithms in matters of ethics, public policy, and social inequality, and many linked their college experiences to issues of algorithmic justice. One student, who had learned in the focus group about personalization efforts by learning management systems (LMS) companies, saw these changes looming large:

Without knowledge, there's no way you can create change, so it's like, you can't fix climate change, if you don't understand where it comes from, or you don't know what exists, which I think is a big problem, and now I think especially of Canvas, I had no idea that LMS was tracking my information for other uses. We're not seeing these things here, so people don't care about them, because they're not seeing it in the media, they're not seeing it in their schools, or wherever we get our information from, it therefore leaves private industries to take over without anyone else knowing.

This understanding of the interplay between individual and societal impacts was evident across the focus groups. Students often expressed confidence in their own ability to work with or around algorithms, having witnessed their rise as a self-described 'pivot generation,' but they were worried about the wider implications for others (see sidebar, "The pivot generation").

When the age of algorithms is in the classroom

Though few faculty members had ready answers for our questions about the classroom, there were some noteworthy exceptions. One social science professor described how he tied social theory to the ways students present themselves online.

They very quickly see how the whole ecosystem is designed to convey not a realistic version of themselves, but an idealized one that allows them to perform a certain kind of identity. And then that links to social movements, and current events, and how we are framing current events through these platforms for consumption, but also to display our own selves as part of a group or another.

Though it was particularly relevant to his discipline, this instructor went on to argue that students are "eager to engage in that kind of inquiry, it's crucial for any informed citizen of the world, and every student should be learning this." Another faculty member frequently demonstrated how tracking can be revealed on websites to open discussion of how large corporations collect and use information. "I share that with every class that I teach."

Similarly, an instructor asked students to think about how wearable technology like FitBit made their personal health data valuable in a variety of unforeseen ways, saying "perhaps the most profound issue of our time is understanding how science and technology in society interact with one another, and it clearly interests students at all grade levels. We're crazy if we don't address it."

Other faculty members raised the topic more circumspectly. Courses that taught students quantitative understanding provided an opening. As one instructor said, "I can't have a data source, and not talk about biases and heuristics." Students also drew connections between algorithms and critical approaches to data; learning about confirmation bias, statistical modeling, and

questioning the sources of data contributed to their ability to understand the ways algorithms shape what they see, even if algorithms were not addressed directly in class. The well-being aspects of technology also provided openings:

We talk about mental health and the effect of living online, and in that sense it kind of comes up. We don't get into specifics about different platforms, but I'd gladly have that conversation with them.

Though a minority of faculty members had found ways to incorporate algorithmic literacy into their courses, several thought it was important for their institution and speculated that the solution was to develop new interdisciplinary courses. One faculty member suggested that the best way to teach about the intersection between technology and society was to bring humanities professors into the discussion since they are accustomed to reasoning through ethical questions. Another faculty member, who teaches an information literacy course, argued that the power of algorithms should be repeatedly addressed throughout the curriculum as a crucial part of understanding the role of information in society. A single "vaccination" approach might be counterproductive.

Students often don't think about how their cell phones track so much of their lives and how the internet tracks so much of their lives in turn, and what this could mean for them as individuals and for society as a whole. So just by bringing that back into the forefront of their mind, they're often very surprised. When it's just brought up once, maybe twice, it's something that you can easily push aside, because it's a scary thought. And if you only think about it once and then you get kind of freaked out, you maybe just want to never think about it again. So it needs to be reinforced over and over again.

All in all though, faculty in our interviews were divided: Though nearly all expressed great concern about the effect of algorithms on our information environment, only a few embraced the challenge of incorporating discussion of algorithms in their courses. Others were hopeful about adding something new to the curriculum, but the majority still believed that their current curricula about critical thinking, and encouraging students to use peer-reviewed sources rather than internet-based sources were applicable and sufficient. As one faculty member pointed out, however, there's no guarantee those lessons have lasting value for students: "Are they going to learn the lesson so well that they will take it with them when they leave and apply it to everything they see for the rest of their lives? That is the challenge."



The “pivot generation”

Throughout the focus group sessions, students expressed concerns about the ability of people older than themselves to navigate systems designed for algorithmic attention and persuasion. As one student said, “My grandfather says ‘Oh, man, Facebook is so addicting.’ I’m like, ‘Yeah, because it’s designed to keep you there, like heroin.’” Another student offered an incisive reflection, “Everyone was so focused on making sure that kids learned that they forgot they also needed to teach grandparents.”

Students were even more worried about the effect of technology on younger people growing up with tablets and phones, especially related to privacy and wellness.* As one student explained:

Now that kids are like learning with iPads and all this new technology from the first day of school, it’s important to make them aware of all the things that are going on behind the scenes, like how they’re personalizing and using all the data from it. Our generation is kind of different, because we’ve learned how to do it right when it came out, and so we are more aware of what’s going on. But if you grew up completely with

all this technology and all that, then you would just have no idea about potentially some of the negative effects that it could have.

Faculty also differentiated the news awareness of this group of students from others they have taught in recent years. As one noted, “They all know about it already, there’s a sense that they have a set of shared knowledge, because they are very up to the minute. They are aware of a difference in terms of how quickly they get information, and they’re aware that it spreads differently through their generation than it does through my generation.” A student echoed this perspective, but worried that these changes came with a cost:

When I listen to younger kids, like kids that are in middle school and high school, they’re talking about climate change, and bias in the media and corruption in politics. And I’m, like, all I cared about when I was 13 was whether my mom would let me get like a bigger pen; what I wanted was moon boots.

While there are many valid criticisms of categorizing people by birth cohort,[†] students in this study characterized *themselves* in generational terms as “the pivot generation.” This is in and of itself unique from prior PIL studies with college students. This sample of students identified as members of a distinct group who came of age at a pivotal moment in the history of technology. Or, as one student put it:

Since we were raised at least for a period of time without this omnipresent influence of social media, we had more of a choice to join the world of social media than the next generation. And I think that we have a lot more perspective on it than the other generations will have. ■

* Some research suggests young children are fairly sophisticated about testing the limits of technology, though many questions remain. See Tanya Basu (6 December 2019), “Why kids don’t trust Alexa,” *MIT Technology Review*, <https://www.technologyreview.com/s/614863/why-kids-dont-trust-alexa>; Judith H. Danovich (2019), “Growing up with Google: How children’s understanding and use of internet-based devices relates to cognitive development,” *Human Behavior and Emerging Technologies* 2, 81-90, DOI: <https://doi.org/10.1002/hbe2.142>

† See, for example, Mark Bullen, Tannis Morgan, and Adnan Qayyum (2011), “Digital learners in higher education: Generation is not the issue,” *Canadian Journal of Learning and Technology* 37(1), <https://www.learntechlib.org/p/42755/>; Ellen Johanna Helsper, and Rebecca Eynon (2013), “Digital natives: Where is the evidence?” *British Educational Research Journal* 36(3), 503-520, DOI: <https://doi.org/10.1080/01411920902989227>

Part Three:

Conclusion and Recommendations

Our report is one of the first multi-institutional investigations into college students' awareness of and concerns about how algorithms are shaping the news and information they receive from internet giants, such as Google, Amazon, YouTube, Instagram, and Facebook. Qualitative data were collected from students participating in one of 16 focus groups, and 37 faculty telephone interviews, at eight U.S. colleges and universities. Together, these discussions examined how students were experiencing dramatic changes brought on by algorithms, while exploring the extent to which the issues of personalization, privacy protection, machine learning, and AI have entered the classroom.

This research was conducted in the wake of coverage of the Cambridge Analytica scandal, one of the most-read news stories in 2018.⁷⁷ Powerful algorithms had trawled through 87 million people's Facebook profiles to sway voters during two hotly contested campaigns: support for Brexit in the U.K. and the presidential election of Donald Trump in the U.S. Since then, algorithms have clearly entered the public conversation, and students in our study, like so many around them, were frustrated with the opaque lines of coding trying to influence their interactions on popular websites.

Most students recognized that as information has become ubiquitous, the hidden levers that personalize results and nudge us toward selected information often camouflage complexity behind the appearance of simplicity and efficiency. Moreover, many, though not all, were aware that data-driven algorithms, if unexamined and unchallenged, could threaten

representative democracy and the cultivation of informed and engaged communities.

Importantly, students were both resigned to and indignant about algorithmic-driven ads, news, and information. Yet, many found sites like Google, YouTube, Instagram, and Facebook too useful to abandon. Many seemed resigned to the powers of an unregulated media environment, but were willing to engage with the platforms to exert their agency and protect their privacy. Their concerns were often accompanied by a sense of impotence, and for some, nihilistic dread. While some students worried about the “creepiness” of algorithms that eavesdrop on their offline conversations to try to sell them a product, others had concerns about the real-life consequences of automated decision-making systems that reinforce societal inequalities.

Faculty in our interviews often expressed frustration and powerlessness with ubiquitous algorithmic systems that affect society. They lamented the “loss of a common culture,” and compromised privacy without accountability to the public. Their response was largely to stick to a narrow set of information sources, like *The New York Times* or NPR, and avoid social media platforms altogether.

A wide and concerning gap

More than anything, findings from this study identify a wide gap between students and faculty as algorithms, put to both good and bad use, are changing the online information landscape in

⁷⁷ Maggie Adams, Ari Isaacman Bevacqua, and Anna Dubenko (19 December 2018), “The most-read *New York Times* stories of 2018,” Story #52: Matthew Rosenberg, Nicholas Confessore, and Carole Cadwalladr, “How Trump consultants exploited the Facebook data of millions,” (17 March 2018). This story about Cambridge Analytica ignited harsh criticism from lawmakers about Facebook's business dealings and use of algorithms. <https://www.nytimes.com/interactive/2018/12/19/reader-center/top-stories.html>. See also, Carol Cadwalladr (April 2019), “Facebook's role in Brexit — and the threat to democracy,” TED Talk, https://www.ted.com/talks/carole_cadwalladr_facebook_s_role_in_brexit_and_the_threat_to_democracy

opaque and unknown ways. In a major finding to emerge from this study, students were far more eager than their faculty to fight back in practical ways against algorithmic control with strategies they learned from friends and peers, but not in class. They realize that it takes reading across different content producers to get the full story, and that this takes time and effort. They know that “choice” is a tricky prospect with online news, so some have figured out how to use social media sites, like Twitter and Reddit, as their “news editors” to help them exert a little control and break out of filter bubbles.

While students’ willingness to contest the workings of these internet giants is encouraging, our findings suggest the age of algorithms demands that teaching strategies be reconsidered as we redefine information literacy. Students should not have to learn these critical information skills on their own. Nor should it be assumed that all of their strategies are necessarily effective.

In our focus groups and interviews, we found a troubling trend aligned with what we already know about students and their information practices from a decade of PIL research studies. That is, the critical work of understanding the torrent of information flowing through a variety of channels, from social media to commercial search engines, is rarely considered in assignments and classroom discussions.

It was surprising to discover how rarely current information systems, and the social and economic conditions that shape and influence their design, were discussed in the classroom. The persistence of this static approach to information, which fails to acknowledge how the world has changed in the 20 years since Google began capturing and exploiting individuals’ digital trails, has powerful consequences. While our exploratory study sample was small and the methods qualitative, these findings warrant further investigation by future researchers.

Nonetheless, it must be acknowledged that the information environment our students inhabit is not a cloister of scholarly knowledge. It more closely resembles an overgrown jungle where every resource must be tested for toxicity, and where students are stalked relentlessly, their data harvested as fodder for unknowable uses.

We do a disservice to our students and to society by confining research assignments and information literacy instruction efforts to the walled garden of peer-reviewed scholarship, where truth is plucked from well-pruned sources and carefully packaged for instructors following explicit instructions. When students are given so limited a range of exploration that they graduate feeling ill-equipped to ask their own questions, higher education has failed them.

Recommendations

These recommendations are provided for key stakeholders — educators, librarians, administrators, and journalists — involved in promoting truth and knowledge in the “post-truth” era. They are grounded in findings from this study, lessons learned from a decade of PIL research, and from discussions with a cross-disciplinary panel of experts convened at the Harvard Graduate School of Education in November 2019.⁷⁸ Above all, they are aimed at enhancing algorithmic literacy.

For the large part, they make students partners in addressing algorithmic education in order to promote widespread awareness of algorithms and strategies that may preserve human agency and autonomy. Striking a balance between being idealistic and practical, they build on what teachers, librarians, and journalists already do to advance public understanding and engagement with information in a fast-changing world.

⁷⁸ Note their comments found on p. 37.

» ***Recommendation 1: Use peer-to-peer learning to nurture personal agency and advance campus-wide learning.***

Students in our focus groups almost always identified their peers as knowledge sources about algorithms. This self-described “pivot generation” felt they were better positioned to question and counter personalized information flows than their older family members and instructors. They learned how systems worked through observation and comparing notes and workarounds with their friends and younger relatives.

In other cases, they connected this personal experience to broader classroom discussions of social justice and politics, bringing a sophisticated critical lens to bear on complex issues related to the flow of information. In stark contrast, the faculty we interviewed for this study were concerned and curious about how algorithmic systems are influencing public life, but they rarely felt personally prepared to broach these issues in their classrooms.

Together, these findings point to the potential of “students as partners” initiatives to raise awareness of algorithms on campus. The emerging literature on developing partnerships between students and faculty, staff, and administrators provides models for practice that actively and authentically involve students in curriculum and service design, and in research and teaching.⁷⁹ Campus teaching and learning centers often have programs already in place to support these partnerships.

In this scenario, educators’ disciplinary understanding of information and society could be combined with students’ knowledge and creative

practices to develop a curriculum that meets the needs of students as citizens, not just as scholars. Rather than advocating algorithmic abstinence, as many of the faculty we talked to might choose to do, this could lead students to examine how algorithms affect their lives. It could lead them to consider when and how they might respond, from employing defensive tactics to pushing for social or legislative change.

At the next level, co-learning positions students as peer teachers. Peer-to-peer teaching is well established in college classrooms including its use to foster digital literacy.⁸⁰ Formalizing the everyday habits of learning students already have while leveraging shared experiences, concerns, and language may be an effective way to increase algorithmic literacy on campus.

This teaching role could be further supported and extended beyond the classroom with students providing assistance for learners, including faculty and staff, as they develop digital and algorithmic literacy. To highlight students’ ownership of their knowledge, a skill-share session could be organized on campus with stakeholders from student affairs, academic departments, IT, and the library in which students lead conversations about the social implications of data-driven decision systems and provide hands-on training on tools and strategies, such as surveillance self-defense.

As institutions examine their own use of personal data, particularly in learning analytics, there is also a role for students to play in designing for transparency,⁸¹ and such efforts could open campus-wide discussions about the intersection of algorithms and agency. At the very least they might

79 Alison Cook-Sather, Melanie Bahti, and Anita Ntem (2019), *Pedagogical partnerships: A how-to guide for faculty, students, and academic developers in higher education*, Elon University Center for Engaged Learning, <https://www.centerforengagedlearning.org/books/pedagogical-partnerships/>

80 See for example, Julia Terry, Alyson Davies, Catherine Williams, Sarah Tait, and Louise Condon (2019), “Improving the digital literacy competence of nursing and midwifery students: A qualitative study of the experiences of NICE student champions,” *Nurse Education in Practice* 34, 192-198, DOI: <https://doi.org/10.1016/j.nepr.2018.11.016>; and the ePioneers program at Oxford Brookes University detailed here <https://jjscinfonetcasestudies.pbworks.com/w/page/74349344/Digital%20Literacies%20at%20Oxford%20Brookes%20University>

81 Mollie Dollinger and Jason Lodge (2019), “What learning analytics can learn from students as partners,” *Educational Media International* 56(3), 1-15. DOI: <https://doi.org/10.1080/09523987.2019.1669883>; Christine Broughan, and Paul Prinsloo (2019), “(Re)centring students in learning analytics: In conversation with Paulo Freire,” *Assessment & Evaluation in Higher Education* (2019): 1-12, DOI: <https://doi.org/10.1080/02602938.2019.1679716>

spark investigations by student journalists.⁸²

The concern students showed for the naiveté of those with less awareness of the media environment could inform peer learning outreach work beyond the institution to the wider community. As communication studies and political science departments develop deliberative democracy programs,⁸³ students could lead discussions about the social and political ramifications of algorithms and how they influence our understanding of current events, especially relevant during a heated political season. (This peer learning model could also be folded into school and adult learning partnerships as sketched out in Recommendation 2.)

Reconceiving of students as partners and incorporating co-learning into the curriculum will require some thoughtful realignment of roles, while calling for vulnerability and trust.⁸⁴ Teachers need to be willing to ask and welcome questions they cannot answer. Students must take responsibility for developing and sharing their knowledge and listening to one another.⁸⁵ Both parties must be willing to set aside the reassuring familiarity of hierarchies of power and information and, instead, encourage curiosity and tolerate ambiguity, even if the result of such learning is difficult to predict and entails risk for both students and the faculty who must, ultimately, assign grades.

Finally, we cannot leave it solely to students to be change agents on campus; librarians and educators need to address their own knowledge gaps. Instructors need to develop a greater understanding of how algorithms affect their own teaching and research, and shape the lives of their students. Librarians could take the lead on campus, forming communities of interest among faculty, identifying campus experts, sponsoring an “algo book club” and using instructional partnerships to help instructors integrate algorithmic literacy into their courses. To build their own knowledge, librarians could form a journal club to share readings, add a “skill share” or “what’s new” technology component to regular staff meetings, and strategize how the library can support learning about the age of algorithms through programming, services, interdepartmental initiatives, and the library’s instruction program.⁸⁶

» ***Recommendation 2: The K-20 student learning experience must be interdisciplinary, holistic, and integrated.***

Students in this study described their exposure to information literacy and critical thinking from elementary school through college as scattered, inadequate, and disconnected. Critical thinking practices that focus on closely analyzing texts can be valuable, but must be accompanied by a nimble set of evaluative strategies for sorting through new information on the fly to cope with

82 Op. cit. Zak Vescera 2019.

83 Nancy L. Thomas and J. Kyle Upchurch (2018), “Strengthening democracy by design: Challenges and opportunities,” *Journal of Public Deliberation* 14(2), <https://www.publicdeliberation.net/jpd/vol14/iss2/art9>; Hayley J. Cole. (2013), “Teaching, practicing, and performing deliberative democracy in the classroom,” *Journal of Public Deliberation* 9(2), <https://www.publicdeliberation.net/jpd/vol9/iss2/art10>

84 Kelly E. Matthews, Lucy Mercer-Mapstone, Sam Lucie Dvorakova, Anita Acai, Alison Cook-Sather, Peter Felten, Mick Healey, Ruth L. Healey, and Elizabeth Marquis (2019), “Enhancing outcomes and reducing inhibitors to the engagement of students and staff in learning and teaching partnerships: Implications for academic development,” *International Journal for Academic Development* 24(3), 246-259, DOI: <https://doi.org/10.1080/1360144X.2018.1545233>; Abbi Flint (2015), “Students and staff as partners in innovation and change,” *The Journal of Educational Innovation, Partnership and Change* 1(1), <https://journals.studentengagement.org.uk/index.php/studentchangeagents/article/view/218>

85 Hayley Burke (2013), “Legitimizing student expertise in student-faculty partnerships,” *Teaching and Learning Together in Higher Education* 10, <http://repository.brynmawr.edu/tlthe/vol1/iss10/6>

86 For examples of a range of peer learning initiatives around technology, see Elizabeth Bennett and Susan Folley (2015), *D4 Strategic Project: Developing Staff Digital Literacies. External Scoping Report*. University of Huddersfield, <http://eprints.hud.ac.uk/id/eprint/26265/1/BennettExternal.pdf>

the volume of choices we face in a world saturated with information.⁸⁷ News and information is no longer something we seek, it seeks us through a variety of channels that clamor for attention.

As algorithms continue to have a pervasive presence in our lives, more must be done to make information literacy instruction coherent and holistic throughout K-20. This is especially true in light of concerns we heard from students about the targeting of children by algorithm-driven platforms promoting commercial products. Students need a greater understanding of how news and information work in, and on, society as they flow and are shaped by algorithm-driven and market-influenced intermediaries. Moreover, we should explore how the domains of reading, writing, and quantitative skills connect to efforts across the lives of students to introduce them to media, news, digital, and information literacy. This integrative work will require the formation of alliances at the local and national levels.

At a tactical level, local efforts could start with finding stakeholders on college campuses who have already built bridges to the local schools and into the community. These may be teacher education faculty, coordinators of programs that pair college students with elementary school students, community outreach programs and student participants. Working through existing connections, educators, students, librarians from public, school, and academic sectors, and representatives from local news organizations could be invited to an open-ended discussion and workshop. Such a gathering could audit what students are learning, map connections across the learning experience, identify gaps, and seek ways to continue working together.

Rather than focusing on inventing new curricula, these conversations could engage participants with questions, such as: What are students at

different levels actually learning? How can that learning be improved and scaffolded so there is greater cohesiveness from K-12 to college and beyond? What will it take to give learners of all ages a better grasp of the hallmarks of trustworthy information and how can they learn about the ethical standards undergirding journalism, scholarship, and science? What do they need to know about how news content is produced and disseminated to develop discernment, and ultimately, where warranted, trust in the news? What do they need to know about how algorithm-driven systems work and affect the information they engage with? Considering the impact of algorithms on information systems and our daily lives, what else needs to be folded into existing instruction, and how can educators and the broader public get up to speed?

More strategically, endorsement from local leadership such as a school superintendent, a college president, or a dean of students could kick start conversations by providing a meeting space and lending their imprimatur to the effort. At the same time, national organizations could work on bridging the different cultures and needs of K-12, college, and community educators, developing a shared understanding of what it means to be information literate today. For example, the American Library Association could facilitate a meeting of interested parties from the American Association of School Librarians, the Association of College and Research Libraries, and the Public Library Association to collaborate on a vision for integrating information literacy efforts across the lifespan.

Another configuration would bring together representatives from professional organizations, such as the Society of Professional Journalists, the Association for Education in Journalism and Mass Communication, and the International Society for Technology in Education, and leading scholars in the field of Science and

87 For examples of curricular approaches that emphasize practical evaluation of web sources, see Mike Caulfield's "Check, Please" lesson plans (<http://lessons.checkplease.cc>) and the Stanford History Education Group's Civic Online Reasoning curriculum collections (<https://cor.stanford.edu/>).

Technology Studies. Policy experts at The Knight Foundation, Pew Research, and others in the nonprofit sector have a key role in seeking common ground and develop a blueprint for educating the public about the impact of algorithms on news and information. Library-focused organizations could meet with those in other domains to build connections. The idea would be to better coordinate and update information literacy and related programs, find common ground, and promote public understanding.

These national efforts would be costly, but not as costly as ignorance. Ideally, funding and sponsorship could be sought to design a sustainable cross-disciplinary curriculum toolkit that could collect and curate exemplary learning materials for learners at all levels,⁸⁸ as well as launching a newsletter to update stakeholders on what's new at the intersection of information, technology, and society.

» ***Recommendation 3: News outlets must expand algorithm coverage, while being transparent about their own practices.***

While participants in this study could see the personal effects of algorithms in the news and advertisements served up to them, it takes the contextualization of solid reporting to demonstrate that these are part of larger patterns with social consequences.⁸⁹ Reporting that includes practical tactics — defensive practices like those we heard about from students — could help counter the narrative of helpless resignation we heard in our student focus groups and faculty interviews.⁹⁰ New tools make it easier for journalists to track the trackers and investigate the algorithms.⁹¹ “Open Sourced,” an initiative from Vox, is a promising development, looking beyond hype and hysteria to focus on “explaining the risks and benefits when it comes to AI and digital privacy so you can make informed decisions.”⁹²

Greater public awareness may shape what companies do and what policies governments enact. There is evidence that there are deep concerns with algorithmic personalization, particularly when it comes to news⁹³ and that this backlash is having an effect on data-gathering efforts and corporate strategy.⁹⁴ Public pressure may lead to improved access to information about

88 See for example, Brett Gaylor's 2015 project “Do Not Track” <https://donottrack-doc.com>, an award-winning, interactive, personalized documentary.

89 See for example, Brandi Geurkink (5 December 2019), “Congratulations YouTube... Now show your work,” *Mozilla Foundation*, <https://foundation.mozilla.org/en/blog/congratulations-youtube-now-show-your-work/>

90 As examples, see Stuart A. Thompson and Charlie Warzel (20 December 2019), “How to track President Trump,” *The New York Times*, <https://www.nytimes.com/interactive/2019/12/20/opinion/location-data-national-security.html>; Geoffrey A. Fowler (27 June 2019), “Help desk: How to fight the spies in your Chrome browser,” *Washington Post*, <https://www.washingtonpost.com/technology/2019/06/27/help-desk-how-fight-spies-your-chrome-browser>

91 See for example Twitter Trails, <http://twittertrails.com/>, the Princeton IoT Inspector, <https://iot-inspector.princeton.edu/>, and Tracking Exposed, <https://tracking.exposed/>

92 Samantha Oltman and Joss Fong (10 December 2019), “Open Sourced: The hidden consequences of tech revealed,” *Vox*, <https://www.vox.com/recode/2019/12/10/20991304/open-sourced-the-hidden-consequences-of-tech-revealed>

93 Nash 2019, op.cit.; Smith 2018, op.cit.; Elisa Shearer and Elizabeth Grieco (2 October 2019), “Americans are wary of the role of social media sites in delivering news,” Pew Research Center, <https://www.journalism.org/2019/10/02/americans-are-wary-of-the-role-social-media-sites-play-in-delivering-the-news/>; Jihii 2014, op.cit.

94 Gartner Group (2 December 2019), “Gartner predicts 80% of marketers will abandon personalization efforts by 2025,” <https://www.gartner.com/en/newsroom/press-releases/2019-12-02-gartner-predicts-80--of-marketers-will-abandon-person>; Sara Fischer (19 November 2019), “New York Times dropping most social media trackers,” <https://www.axios.com/new-york-times-social-media-trackers-47f547d5-a241-424a-a398-12344c78ac32.html>

proprietary algorithms that would facilitate more third-party monitoring by experts and journalists; so far the internet giants have shown little interest in facilitating this.⁹⁵

As the news industry has come under greater scrutiny, journalists have been called upon to help people distinguish objective news coverage from misinformation and outright lies. These demands of journalistic duty to democracy have become greater as algorithms permeate daily life. Journalists must use their platforms to shed light on how algorithms work and de-mystify them in clear language. This requires developing deeper expertise of their own and working with a wider array of academic and technical experts to deepen their investigations.⁹⁶

Too often, journalism plays into dangerous anthropomorphization, granting more agency and power to the systems than they actually have. We need reporting that questions both good and bad uses of algorithms, including their benefits and harms, and tells the stories of impact on individuals within the context of wider society.⁹⁷ This need is real: In our focus groups and faculty interviews, it became clear that many people, no matter how educated, did not understand the way algorithms shape the flow of information.⁹⁸

As editors and journalists at professional news outlets weigh the benefits of integrating algorithms into their business and reporting models, they have a responsibility to be transparent and ethical about their own practices, too. Algorithm-based tools have also become integral to the production of news, from story generation to source identification.⁹⁹ As journalists learn to use these resources, they need to understand the limitations and ethical implications of relying on automated filtering and decision making.¹⁰⁰

While audiences for news may be increasingly aware that algorithms shape what they see, they may not know that similar sets of filters are determining what alerts the journalist to a story, or directs the sources reporters contact or angles they develop.¹⁰¹ This confirms our findings about the general lack of knowledge our students — and others — have about the way that news is produced. Media organizations need to be transparent about how they're using these tools to create their content.

Greater transparency, however, is even more urgently required around the use of algorithms by news outlets to target news and advertising.¹⁰² Students and faculty in this study expressed

95 Jeremy B. Merrill and Ariana Tobin (28 January 2019), "Facebook moves to block ad transparency tools — including ours," *ProPublica*, <https://www.propublica.org/article/facebook-blocks-ad-transparency-tools>; Craig Silverman (22 August 2019), "Facebook said it would give detailed data to academics. They're still waiting," *Buzzfeed News*, <https://www.buzzfeednews.com/article/craigsilverman/slow-facebook>

96 Ren LaForme (15 March 2018), "The next big thing in journalism might be algorithm reporters," *Poynter*, <https://www.poynter.org/tech-tools/2018/the-next-big-thing-in-journalism-might-be-algorithm-reporters/>

97 Francesco Marconi, Till Daldrup, and Rajiv Pant (14 February, 2019), "Acing the algorithm beat, journalism's next frontier," *NiemanLab*, <https://www.niemanlab.org/2019/02/acing-the-algorithmic-beat-journalisms-next-frontier/>

98 Op cit. John Wihbey (2019).

99 Nicholas Diakopoulos (November 2018), "An algorithmic nose for news," *Columbia Journalism Review*, https://www.cjr.org/tow_center/an-algorithmic-nose-for-news.php

100 Paul Cheung (21 November 2019), "Journalism's superfood: AI?" <https://knightfoundation.org/articles/journalisms-superfood-ai/>; Op. cit. Nicholas Diakopoulos (November 2018); Amy Batt and Jacob Granger (11 November 2019), "Artificial Intelligence is not the future - it is happening right now," <https://www.journalism.co.uk/news/artificial-intelligence-is-not-the-future-it-is-happening-right-now/s2/a747107/>

101 Op cit. John Wihbey (2019), see especially Chapter 6, "Data, Artificial Intelligence, and the News Future."

102 Cheung (21 November 2019), op.cit.; Sophia Ignatidou (December 2019), "AI Driven personalization in digital media: political and societal implications," <https://www.chathamhouse.org/sites/default/files/021219%20AI-driven%20Personalization%20in%20Digital%20Media%20final%20WEB.pdf>; United States, Cong. Senate, Filter Bubble Transparency Act. S. 2763, Washington: GPO, 2019, https://www.thune.senate.gov/public/_cache/files/c3a43550-7c36-4f77-b05c-d2275c0d568c/CE3DDB84DDB9284CC6D372833D039A20.filter-bubble-final.pdf; Max Z. Van Drunen, Natali Helberger, and Mariella Bastian (2019), "Know your algorithm: What media organizations need to explain to their users about news personalization," *International Data Privacy Law*, DOI: <https://doi.org/10.1093/idpl/izp2011>

concern about the quality of news available to them. Many stated that algorithm-driven personalization increased their distrust of news and raised additional issues, particularly the potential for news silos. Some media organizations are responding to this concern by making their personalization policies more openly available,¹⁰³ undertaking deeper studies to ensure a responsible approach,¹⁰⁴ and in some cases, moving away from data sharing with other platforms.¹⁰⁵ Still, on most news sites, transparency information is difficult to find, hard to read, and incomplete.¹⁰⁶ To regain the trust of their audiences, media organizations need to be much clearer about what information they collect, how they use it and with whom they share it.

» **Recommendation 4: Learning about algorithmic justice supports education for democracy.**

Despite their aura of sophisticated cynicism, students in our focus groups often became energized when discussing the impact of

algorithms on equality, status, inclusion, and opportunities. As one student noted, “It worries me if it’s systemically allowing certain groups to succeed over others.” Another observed “our moral compass seems to be broken online.” Comments like these present a rich opportunity for engaged learning and civic engagement. Though students in our study expressed helplessness in the face of powerful corporations, they became motivated to challenge them as they learned more. This faultline between the perception of helplessness and a desire to create change is a productive site of emotional friction that opens opportunities to engage in “education for democracy.”¹⁰⁷

News reports remind us daily that there is work to be done: Cities debate the ethics of facial recognition systems¹⁰⁸ and residents question the use of doorbell surveillance to monitor neighborhoods;¹⁰⁹ legislators wrestle with regulating data-gathering;¹¹⁰ extremists and pedophiles use popular platforms to groom the vulnerable;¹¹¹ librarians and educators raise concerns about commercial products that harvest data from students;¹¹² software engineers question

103 “Personalization,” (14 May 2018), *The New York Times* <https://help.nytimes.com/hc/en-us/articles/360003965994-Personalization>

104 The British Broadcasting Corporation (2018), “Responsible machine learning in the public interest: Developing machine learning and data-enabled technology in a responsible way that upholds BBC values,” <https://www.bbc.co.uk/rd/projects/responsible-machine-learning>

105 Fischer 2019, op.cit.

106 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://timlibert.me/pdf/LIBERT_BINNS-2019-GOOD_NEWS.pdf

107 Philosopher John Dewey wrote about a progressive approach to education over a century ago in *Democracy and Education* (1916) and other works which continue to influence educators. See for example, Tomas Englund (2000), “Rethinking democracy and education: towards an education of deliberative citizens,” *Journal of Curriculum Studies*, 32(2), 305–313, DOI: <https://doi.org/10.1080/002202700182772>

108 Blake Montgomery (31 July 2019), “Facial recognition bans: Coming soon to a city near you,” *The Daily Beast*, <https://www.thedailybeast.com/facial-recognition-bans-coming-soon-to-a-city-near-you>

109 Caroline Haskins (3 December 2019), “How Ring went from ‘Shark Tank’ reject to American’s scariest surveillance company,” *Vice*, https://www.vice.com/en_us/article/zmjp53/how-ring-went-from-shark-tank-reject-to-americas-scariest-surveillance-company

110 Dominic Rushe (26 November 2019), “Democrats propose sweeping new online privacy laws to rein in tech giants,” *The Guardian*, <https://www.theguardian.com/world/2019/nov/26/democrats-propose-online-privacy-laws>

111 Anonymous. (5 May 2019), “What happened after my 13-year-old son joined the alt-right,” *Washingtonian*, <https://www.washingtonian.com/2019/05/05/what-happened-after-my-13-year-old-son-joined-the-alt-right/>; Nellie Bowles and Michael H. Keller (7 December 2019), “Video games and online chats are ‘hunting grounds’ for sexual predators,” *The New York Times*, <https://nyti.ms/2qx6fmm>

112 Kyle M. L. Jones and Barbara Fister (14 October 2019), “Kyle M.L. Jones: The datafied student and the ethics of learning analytics,” [email interview] by Barbara Fister, Project Information Literacy, Smart Talk Interview, no. 32, <https://www.projectinfolit.org/kyle-jones-smart-talk.html>; Rebecca Koenig (17 October 2019), “At Educause, a push to monitor student data is met with concerns about privacy and equity,” *Edsurge*, <https://www.edsurge.com/news/2019-10-17-at-educause-a-push-to-monitor-student-data-is-met-with-concerns-about-privacy-and-equity>; Kyle M.L. Jones (2 July 2019) op. cit.; Data Doubles, <https://datadoubles.org/>; Audrey Watters, Hack Education, <http://hackeducation.com/>

the morality of their work.¹¹³ Every day new issues surface. It is around pressing topics like these that students can break through a sense of helplessness, be emboldened with personal agency to grapple with complex issues, and feel empowered to take on the challenge of promoting algorithmic justice.

At a practical level, individual instructors can look out for stories in the news that link their subject matter to issues of algorithmic justice: How does the digital surveillance of children influence child development? What information could help hospitals follow up with patients without introducing bias? How does microtargeting ads for jobs and housing relate to the history of redlining? Librarians who serve as liaisons to academic departments could support these efforts by creating ongoing curated collections of relevant news stories targeted to specific courses and disciplines, strengthening their own algorithmic literacy while broadening the working definition of information literacy on campus.

By injecting current controversies around the algorithmic systems that influence our lives into their course material, educators can tie their disciplinary knowledge to pressing questions of ethics, fairness, and social justice. As we learned in PIL's 2018 news study, the classroom is an influential incubator for the discussion of news and the interpretation of current events; almost two-thirds of survey respondents had learned about news from faculty discussions during the previous week.¹¹⁴

Librarians have developed programs of their own that can be shared across and beyond the campus. Two programs funded by the Institute of Museum and Library Services develop capacity for librarians to take on the challenges of our digital environment. The Library Freedom Project trains librarians to become local experts on privacy practices who can take their knowledge into their communities.¹¹⁵ The Algorithmic Awareness Project is developing a curriculum, syllabi and software for educating librarians and developing open educational resources.¹¹⁶ Many libraries have stepped up to develop guides and workshops for their communities¹¹⁷ as well as credit-bearing courses and open educational resources.¹¹⁸ More than ever, librarians are bringing issues of social justice and information systems into their teaching, tying digital ethics to information literacy.¹¹⁹

Several research programs are developing interesting resources for algorithmic literacy. Researchers at the University of Amsterdam's Department of Media Studies are developing tools to allow users to compare and reflect on how social media platforms personalize and filter their information, enabling "data activism" — an intervention that could be used to help students understand and conduct research on social media.¹²⁰ The UnBias Project in the U.K. has created "youth juries" to involve students in weighing concerns about algorithms and proposing solutions.¹²¹ Their Fairness Toolkit

113 Arielle Pardes (22 November 2019), Google employees protest to 'fight for the future of tech,'" *Wired*, <https://www.wired.com/story/google-employees-protest-retaliation/>

114 Op. cit. Head, et al. 2018, *How students engage with news: Five takeaways for educators, journalists, and librarians*.

115 Library Freedom Project, <https://libraryfreedom.org/>

116 Jason Clark, *Algorithmic Awareness*, <https://github.com/jasonclark/algorithmic-awareness>

117 For example, a workshop and guide was developed to accompany a public event about mass surveillance (<http://libguides.gustavus.edu/mayday>)

118 A textbook created by Barbara Fister, Rachel Flynn and students in IDS 101 for a course titled *Clickbait, bias, and propaganda in information networks* is a handbook for understanding and evaluating information in a networked environment that includes student-authored chapters, <https://mlpp.pressbooks.pub/informationnetworks/>

119 For example the Library and Information Technology Association webinar, "Engaging with algorithm bias: How librarians can meet information literacy needs of computer science and engineering students" on 16 December 2019. <http://www.ala.org/lita/engaging-algorithm-bias-how-librarians-can-meet-information-literacy-needs-computer-science-and>

120 *Algorithms Exposed*, <https://algorithms.exposed/>

121 UnBias Project, <https://unbias.wp.horizon.ac.uk/>

can be downloaded and adapted in classrooms K-20 to encourage civic learning and action.¹²² Student-centered activities like these promote a “collective approach to imagining the future as a contrast to the individual atomizing effect that such technologies often cause.” At the same time, a team of researchers at MIT’s Media Lab is working on an “Algorithmic Justice” program to help students understand how to navigate algorithmic systems from the inside out, by designing them from scratch using a series of low-cost “unplugged” activities.¹²³

Instruction in racial literacy¹²⁴ intersects powerfully with algorithmic literacy. The computer science curriculum must include a component of racial literacy to ensure future coders are cognizant of the ethical considerations they must bear in mind as they design new systems.¹²⁵ History courses could develop a unit on how early computing efforts melded with eugenics contributed to the Holocaust, and link that history to contemporary controversies about the rise of online extremism. One programmatic idea is to pair a scholar of racism with a computer scientist to lead public discussions of how social media platforms and the communities that use them could address online harassment, race-based targeting, and the spread of extremist propaganda.

Opportunities to introduce learning about algorithmic justice can be found throughout the curriculum. At the college level, a group of interested faculty from across the curriculum could conduct a curriculum-mapping project to engage students in the social dimensions of the present moment. As students proceed through their education, they could encounter the rich intersections of the humanistic, social, technical, and quantitative aspects of algorithms that

could help them connect their lives with social trends, their learning across disciplines, and their personal lives with broader issues of social justice.

We are facing a global epistemological crisis. People no longer know what to believe or on what grounds we can determine what is true. It is imperative that truth-seeking institutions — education and journalism — take the lead in healing the social fractures that technology has widened. The technical infrastructure that channels and shapes so much of our understanding and social interaction was created in the utopian belief that making information universally available and giving every individual a voice would improve our lives. But as that infrastructure became an engine of surveillance and persuasion, trading in the intimate details of our lives to create sophisticated marketing tools to sell consumer goods and ideas, that utopian ideal has become dystopian. The power of machine learning and artificial intelligence has been unleashed without regulation or informed consent.

It is no wonder both students and faculty in this study felt helpless and anxious about the future. These recommendations show a path forward. As students claim their authority as learners, as algorithmic literacy is woven into education across the curriculum and students’ life spans, and as journalists give the public tools to understand this epistemological crisis, we will be better prepared to tackle both the unchecked power of algorithms as well the social problems they expose and exacerbate. This education for democracy — both formal and beyond — can empower us to reclaim our role in shaping the future.

122 Fairness Toolkit, UnBias Project, <https://unbias.wp.horizon.ac.uk/fairness-toolkit/>

123 Michelle Ma (13 May 2019), “The future of everything: How to teach kids about AI,” *The Wall Street Journal*, <https://www.wsj.com/articles/how-to-teach-kids-about-ai-11557759541>

124 Winona Guo (27 May, 2019), “Toward a culture of racial literacy,” *Harvard Political Review*, <https://harvardpolitics.com/harvard/toward-a-culture/>

125 Op. cit. Jessie Daniels et al 2019.

Leading Thinkers

To explore the implications of this study's findings, we brought together a select group of leading thinkers in education, libraries, media research, journalism, and technology for a half-day workshop at Harvard Graduate School of Education (HGSE) on November 7, 2019. PIL co-researchers presented early findings from the Algo Study, and participants were invited to share and discuss the implications of our findings and ideas, experiences, solutions, challenges, and puzzling aspects about life in the age of algorithms. Further, they were asked to use their expertise to inform different solutions to issues “truth workers” — educators, librarians, and journalists — face when helping young adults understand the complexities and challenges of our current information environment.

As a final step in the workshop, each participant was asked to submit a reflection responding to this prompt: What do people need to know most about the impact of algorithms, and where and how can we effect change?

Steven Braun

Visiting Assistant Professor, Northeastern University
<https://www.stevengbraun.com>

Every day, we interact with systems that have been engineered to carefully control our access to information, often without realizing that access has been rearranged and modulated by the same systems we trust to read the news, connect with friends and family, or do our work. The rapid proliferation of these algorithm-driven platforms — encouraged under the guise of making our daily lives more efficient, productive, and profitable — has displaced our capacity for doing the critical interpretive work these systems demand, stripping individuals of the responsibility to interpret the facticity of the information these systems give us when that interpretation has been performed by the algorithms themselves.

These algorithms pervade nearly every domain in our lives, and in that ubiquity, we are subconsciously challenged to flex many different literacies — information, data, visual, statistical, and technological literacies, to name a few — to interrogate the intentions of those systems, how they are constructed, and how they control our access to information. Whether we realize it or not, this continuous assault can quickly fatigue our ability to question those systems themselves, tightening the control they have over us even further. In the face of this engineered reality, it can feel easy to become nihilistic about the agency we exercise in responding to these systems as they reconstruct our daily lived experience. Perhaps, the most radical act we can perform, then, is to reject that nihilism and recognize our role as actors in exposing how these systems deprivilege complex, nuanced, and situated knowledges in favor of objective, privileged, and canonical narratives. By reclaiming the responsibility for doing the interpretive work these algorithms have taken from

us, we can become more critical consumers of the information they yield and ultimately encourage a culture in which those systems are designed and consumed in more intentional, ethical ways.

Kasia Chmielinski

*Project Lead and Co-Founder,
Data Nutrition Project*¹²⁶

<https://www.linkedin.com/in/kchmielinski/>

Although the impulse is to believe in the objectivity of the machine, we need to remember that algorithms were built by people. When a human judge makes a decision, it is within our right to respectfully disagree. We understand that people are not infallible and we leverage frameworks like the appeals process to address what we feel are injustices. We should consider algorithmically-determined judgements in the same exact manner; if we blindly accept algorithmically-determined decisions, we are giving up the very important right to appeal and investigate what could be injustices on the part of the algorithm. The algorithm is only reflecting what it has been taught to believe, and often by a very homogenous group of people. Thus, the ability to appeal, alongside the importance of diversifying the pool of people who are building algorithms, can move us along a path towards better artificial intelligence.

Jason A. Clark

*Professor, Head, Special Collections & Archival
Informatics, Montana State University Library*

<http://www.jasonclark.info>

At the risk of wearing the tinfoil hat, I'm suggesting that the first thing people need to know about algorithms and their impact is that many of our everyday decisioning systems are applying algorithmic processes. Insurance coverage decisions, medical analytics, credit card eligibility decisions are increasingly subject to these computational rules. The technological turn in our society, which let me be clear, I have built my career within and love the potential of this turn, makes algorithms an everyday fact. Abstaining from social media or making informed decisions while participating in common technologies will not fully protect you. As for next steps, we have recognized within our *Algorithmic Awareness* project work¹²⁷ that teaching around the primary concepts and introducing essential data literacies that underpin algorithms is a start, but it's not enough. Beyond basic algorithmic literacy, three approaches come to mind: transparent tools, regulation, and technological watermarks. First, we need to start building and introducing tools where transparency about data use and computational decisions is a feature, not a bug (e.g., see tools like the Brave Web browser or the DuckDuckGo search engine). Second, we need more regulation like the General Data Protection Regulation (GDPR) which asserts and confers the "right to explanation" (a right to information about individual decisions made by algorithms).¹²⁸ And finally, we need to consider how technological watermarks such as public/private key fingerprints or an array of byte sequences within files could "watermark" when files have been subjected to algorithmic processes or enhancements (e.g., in the era of deep fakes, this digital authenticity fingerprint will be an essential marker).

¹²⁶ Hilary Ross and Nicole West Bassoff (29 November 2018), "The 'Dataset Nutrition Label Project' tackles dataset health and standards," *Medium*, <https://medium.com/berkman-klein-center/the-dataset-nutrition-label-project-tackles-dataset-health-and-standards-658dc162dffb>

¹²⁷ Jason A. Clark, "RE:Search - Unpacking the algorithms that shape our UX." Deliverables include a teaching curriculum, syllabus, and a software prototype that demonstrates algorithms in action, <https://github.com/jasonclark/algorithmic-awareness>

¹²⁸ Margot E. Kaminski (2019), "The right to explanation, explained," *Berkeley Technology Law Journal* 34(1), DOI: <https://doi.org/10.15779/Z38TD9N83H>

Andrea L. Guzman

*Assistant Professor, Department of Communication,
Northern Illinois University*

<https://andrealguzman.net>

An ongoing challenge regarding algorithms and related technologies, such as artificial intelligence, is demystifying these technologies so that people can act upon knowledge grounded in the reality of the technology instead of its imagined nature. Because algorithms lack a concrete form, people often struggle to clearly articulate what algorithms are and what they do. Conceptualizations of algorithms also may vary from person-to-person. This abstract nature of algorithms contributes to some of the attitudes captured within this study: that algorithms are technological actors that are too difficult to even try understanding or that are beyond the control of the average person. Efforts to help people understand algorithms need to continue moving away from a focus on building awareness of algorithms — people increasingly know about “those things called algorithms” — and toward explaining algorithms in such a way that people have a more consistent conceptualization of what algorithms are, what algorithms do, and — what often is overlooked — what algorithms cannot do.

Key to these efforts is delineating between the role and power of humans in relation to that of the algorithm, so that people understand that algorithms are technologies produced by humans and enacted within human systems. Discourse surrounding algorithms must be careful to not overplay the agency of the algorithm relative to that of people, which often is the result of anthropomorphizing algorithms and their actions. Furthermore, while people need an understanding of algorithms that enables them to take practical actions regarding their personal technology use, such educational efforts should not place the entire onus of taking advantage of or mitigating the effects of algorithms on the individual. There are implications of algorithms, both positive and

negative, that are well outside the control of an individual that may be better addressed within institutions, among community members, or by the government.

Nate Hill

*Executive Director, Metropolitan New York Library
Council (METRO)*

<http://natehill.net/>

Listening to the preliminary findings from the algo study left me reflecting on networked spaces as a medium made for individuals, rather than communities. This is a bit counterintuitive. I think we all expected the opposite. We planned for this utopian, freeing “cyberspace” community, but instead we have been sold a network for mass personalization that ultimately rewards individual vanity over community building. This really clicked when I heard about study participants (students) embracing a certain amount of targeted advertising as long as it benefited them, but then transitioning to calling this “creepy” when they were reminded of the scale of this kind of activity. My takeaway: we need to stop talking about the internet as some kind of inherent good, some kind of human right. This is inaccurate; advancing the interests of individuals above those of the greater community is wrong.

Shortly after the “thinking leaders” session I attended a talk featuring the leader of an excellent organization that is focused on connecting all communities to affordable broadband. The speaker kept on talking about our “transition to a digital society” as some kind of inevitable phenomenon. I know librarians frequently talk about the need for ubiquitous and equitable internet access by citing examples like “people cannot even apply for jobs without going online now.” Have we as a society just succumbed and accepted that efficiency and profit will drive *all* activities into a networked space? What are the ramifications of that? It seems problematic to accept that all activities

will inevitably move online, therefore people all need to be online. Perhaps, at the same time that we focus on connecting everyone, we should talk more about what activities should and shouldn't happen on a network. Should we do all of our reading online? Should the census be digital? If we did disagree with some kind of activity moving online, how would we handle that? So as grumpy as it may sound, I am concluding that human beings need less mediated connection and more human connection.

Alex Hodges

Director of Gutman Library and Faculty, Harvard Graduate School of Education

<http://orcid.org/0000-0003-1712-2816>

As teachers of information literacy, librarians, in conjunction with their teaching partners, absolutely must alert students to better understand their information privacy rights. This concern is especially true as students engage more and more with non-university-supported, third-party products for digital scholarship creation (e.g. citation management, data visualization, multimedia presentation tools, etc.). Our contemporary scholarly communication tools and their connected learning analytics capabilities have broadened the need for society's deeper understanding of digital literacies. We need to have constant dialogue about who owns and has access to individual user data. Such literacies require that learners (or users of various products) understand how to reason for themselves. In addition to these literacies, students also need broader instruction on the information architecture that scaffolds so much of our digital economy — this is where machine learning and algorithms present frontstage in the information literacy constellation. As higher education adopts new models for core curricula and required data science courses, I foresee a ripe future for librarians to expand their teaching and curation roles to advance these additional concerns within information and digital literacy learning.

Momin Malik

Postdoctoral Data Science Research Fellow
Berkman Klein Center for Internet & Society,
Harvard University

<https://www.mominmalik.com/>

A core function of humanities and social science education is to reveal and interrogate the categories, concepts, and logics by which we make sense of and act in the world. Machine learning is reifying one specific logic: that of *statistical correlations*, with no consideration of causality or meaning, a point lost amidst the focus on the “algorithms” that calculate and apply these correlations. The task of higher education will now have to engage with the limits of this logic. To quote William Bruce Cameron on how “not everything that can be counted counts, and not everything that counts can be counted,” the old maxim that “correlation does not imply causation,” and George E. P. Box on how “all models are wrong but some are useful,” we can consider *when, how, for what, and to whom* it is useful to see the world only through correlations between measurable quantities.

For example, “creditworthiness” is an abstract idea, neither the same as future loan repayment, nor of past repayment behavior: we can easily imagine a financially responsible person whose circumstances would nevertheless prevent repayment and are correlated with others as “defaulting.” Should that person be deemed creditworthy or not? Using machine learning to decide forces an answer: The only thing that matters is what correlates with aggregate past behavior. Intention, effort, individuality, and circumstances do not matter, since they do not exist in data. Nor do causal relationships, as studied in econometrics, nor measurement validity, as studied by psychometrics (both statistical fields with their own narrowness and legacy of racism and inequality), nor the possibility of systems being gamed, insofar as they limit the complexity of correlations that can be considered.

Articulating and challenging the frames of “algorithmic” systems will encourage deliberation on the sorts of uses we will accept, and what we will organize to reject.

Panagiotis Takis Metaxas

Professor and Chair, Computer Science, Media Arts & Sciences, Wellesley College

<https://www.wellesley.edu/cs/faculty/metaxas>

What do people need to know most about the impact of algorithms?

People need to understand that “algorithms” are not something new or unusual: It is any process that we follow to achieve a goal (from baking a pie to figuring out how to change a broken tail light on our car). While we casually refer to “algorithms” as the technology that has created the problem, it is really a specific type of algorithms we are worried about: machine learning algorithms that are solving a problem (their output) by looking at the ways that people have already solved it in the past (their input). However, as a famous motto in Computer Science goes, “garbage in? garbage out!” When we give as input to these algorithms data that are problematic, we should not expect that we will get back magically unproblematic solutions.

The problems propagated by these algorithms are mainly related to the way they are used in social media. It is not just a matter of protecting our privacy or keeping our thinking clear of propaganda. It is the fact that the social media companies that use these algorithms do not respect us as individuals. They do not even see us as customers — it is the advertisers that are their customers. They see us as the raw material used to study our behavior so that they can change it to benefit their customers.

Where and how can we effect change?

Education is still the major tool we have to protect ourselves, our cultures, our democracies. It is encouraging that some young adults are aware of the problems created by the algorithms and try to protect themselves, but the challenging issue is to educate the vast majority of our adults, young and old. And we need an education that is reaching people widely and fast. Semi-jokingly, I would say that we need an education that reaches people at the speed of religion.

Eni Mustafaraj

Assistant Professor of Computer Science, Wellesley College

<http://cs.wellesley.edu/~eni/about.html>

OK Computer,¹²⁹ let’s count the ways in which computing devices connected to the internet have rewired human’s media consumption and communication habits in a modern, privileged society. We used to read ink-printed newspapers delivered daily at our doorsteps, whereas now we skim headlines in our favorite online news aggregator on a mobile phone. We used to wander around in department stores’ aisles where new music was sold in packaged CD-ROMs with colorful artwork, whereas now we stream it non-stop on Spotify, often without knowing the artist’s name. Such a list is long. Technologists and venture capitalists want us to believe that this progress is good, desirable, and unstoppable, but we must pause and ask ourselves: who chooses the news headlines we read, and how? Who recommends the next music track to listen, and whose music is left out? The answer is that increasingly algorithms are in charge of these decisions.

But couldn’t we humans make such decisions better ourselves? Why rely on algorithms? A pro-algorithm argument would posit that to include most of humanity in this new modern

¹²⁹ This is the title of the third studio album by the English band Radiohead that came out in 1997.

and interconnected world, we also need speed and scale. An algorithm, which resembles a recipe for cooking a dish, can be executed at light-speed by millions of computers, accomplishing in just a fraction of a second something that would take humans years. If we want more people in the world to have access to the total human knowledge accessible on the Internet, we need algorithms. However, what we need to object against are the values driving the companies that own these algorithms. To effect change, we must collectively advocate for algorithmic-driven information platforms that operate like public libraries: a common good, whose primary goal is not to serve as money-generating machines for share-holders, but to become shared spaces for intellectual and spiritual human flourishing.

Ronald E. Robertson

Doctoral student, Network Science,

Northeastern University

<http://ronalderobertson.com>

The impact of an algorithm is inherently tied to the people who designed it, the information systems it operates on top of, and the people who use it. For example, financial incentives often dictate what algorithms are optimized for, historical data are often biased in ways that perpetuate injustice, and users often have dynamic strategies which guide their behavior. It's important to know that algorithms built upon existing sociotechnical systems, and examining their components in isolation will never tell the whole story. Attempting to do so is like trying to explain tides without considering the moon.

Interdisciplinary research collaborations are therefore essential to understanding the impact of an algorithm. Digital trace data is useful but insufficient. We must also understand who is using them, how they are using them, the information ecosystem they operate in, and how the biases in these dynamic elements interact. However,

interdisciplinary research is hard, and we should expect growing pains. Indeed, computer scientists often miss or neglect important theoretical work, social scientists are often not equipped with the technical skills to gather digital trace data that could ground their theories, and both groups are often critical of one another.

In order to effect change, we will need to bridge this gap — by fostering collaborations, acknowledging the shortcomings of our approaches, and building an appreciation for the value of theoretically-driven mixed methods. As we wait on legislation for protecting users, and on industry collaborations for obtaining data, change will come from independent algorithm audits. Under the incentives of capitalism, and with growing user privacy concerns, we cannot expect or wait upon corporations to cooperate. We must come together to identify the values embedded in their algorithms, spread awareness of their impacts, and develop tools for exposing those impacts and empowering users to overcome them.

Methods

The research findings presented in Part Two of this report are based on qualitative data collected from a random sample of students ($N = 103$) and faculty ($N = 37$) at eight U.S. colleges and universities during Fall 2019. Prior to any data collection for this study, a research protocol was submitted and approved to the Institutional Review Board (IRB) at the Harvard University, where the study was based and at several schools requiring their own IRBs.

Those who voluntarily participated in this study were asked questions in three areas: (1) their awareness and experiences with influential online platforms that use algorithms to shape the flow of information, (2) their perspectives on how information quality may be affected by algorithms, and how they think their own realities and abilities to distinguish truth from falsehoods is affected, and (3) whether and to what extent the impact of algorithms on society was discussed in class.

A script with the five open-ended questions was developed for use with students and faculty.¹³⁰ Demographic data about each of the two sample groups were collected as part of the participation. In both sessions, algorithms were defined as “lines of coding you don’t see that are intentionally used by many online platforms to personalize content to match users likes and dislikes.” News was defined as “information about events happening all around the world.”

Data collection

Between September 9, 2019 and October 9, 2019, a six-person PIL Team collected empirical data from 16 student focus groups and 37 telephone interviews with faculty at eight U.S. colleges and universities. Institutions in the sample were selected for regional diversity, students’

demographic variation, and whether they were located in red or blue states, given the 2016 definition of these voting categories; both a R1 university and a community college were included in the sample (Figure 3 and Table 2).

An email invitation was sent to a randomly selected sample of full-time students on each campus asking them to participate in an hour-long focus group. Two focus groups were held on each campus. A \$20 gift card was offered to all participants who attended all or part of the student focus group. An email invitation was also sent to faculty to participate in a 25- to 30-minute telephone interview. At schools with a sizeable number of teaching faculty, a random sample was used.

Student sample

A total of 103 students from a variety of disciplines participated in one of the two 60-minute focus groups we held on each campus in the library. Groups ranged in size from two to eight students. Qualitative data were collected about students’ use of what we call “internet giants”: sites massive in scale and scope such as Google, YouTube, Twitter, Facebook, and Instagram, that are frequently used by students and faculty for news and information.

As Table 3 (see the end of this section) indicates, the focus group participants were 18 years of age or older and registered as full-time students in their second-year (29%), third-year (31%), fourth-year (28%), or fifth-year or beyond (12%) at the eight colleges in the institutional sample. More than two-thirds of the sample was female (67%). The most common major for participants was social and behavioral studies (18%) and arts and humanities (14%) while far fewer declared computer science (1%) as their majors.

¹³⁰ These scripts are available on the “PIL algorithm project” landing page at https://www.projectinfolit.org/algo_study.html

Figure 3: Institutional sample map

N = 7 U.S. four-year colleges and universities and 1 community college

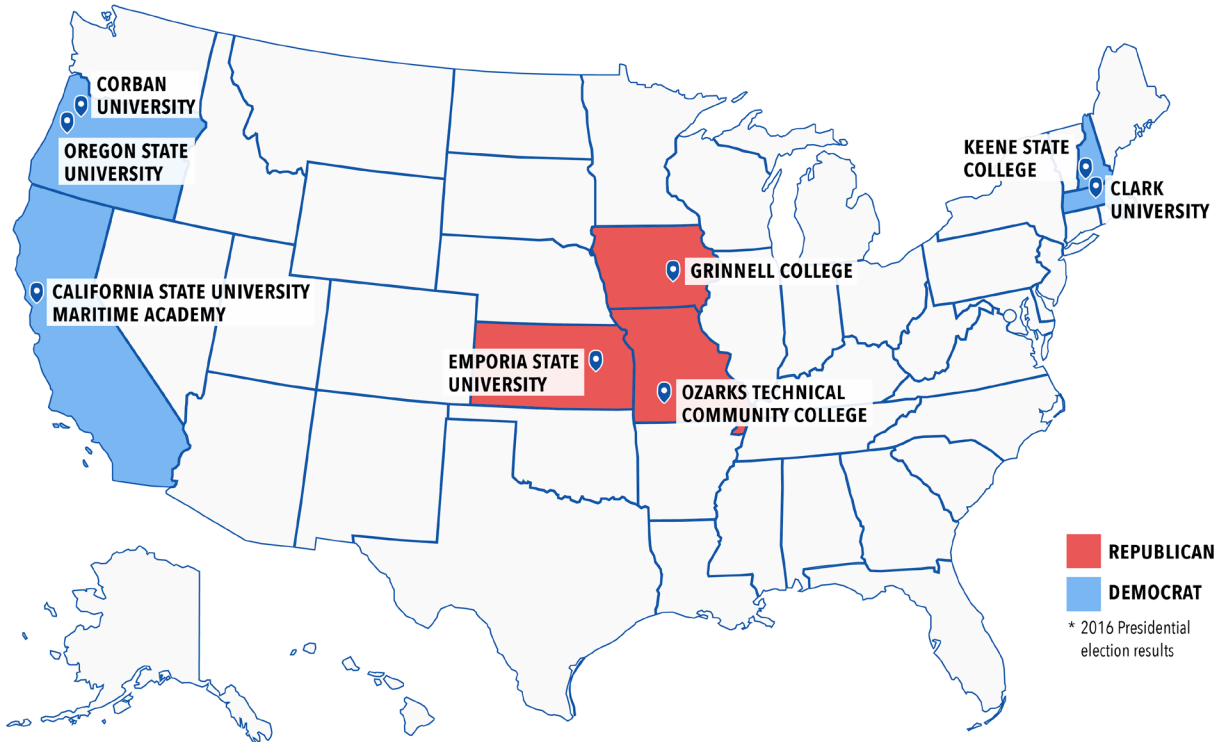


Table 2: Institutional sample

College or University	Location	Type	Focus Groups	Faculty Interviews
California State University Maritime Academy	Vallejo, CA	Public	10	2
Clark University	Worcester, MA	Private	15	8
Corban University	Salem, OR	Private	16	3
Emporia State University	Emporia, KS	Public	12	5
Grinnell College	Grinnell, IA	Private	17	5
Keene State College	Keene, NH	Public	14	4
Oregon State University	Corvallis, OR	Public	8	7
Ozarks Technical Community College	Springfield, MO	Public	11	3

Faculty sample

During the same timeframe as the student focus groups, the PIL Team conducted and recorded telephone interviews with a voluntary sample of faculty at the same schools. The interview script included the same five open-ended questions that were used to collect qualitative data from the student focus groups.

As Table 4 indicates (see the end of this section), the faculty interview sample were 30 years or older than students. Two-fifths of the faculty sample were older than 50-years-old at the eight colleges in the institutional sample (Table 4). Faculty had taught at the institution where they now worked for 1-6 years (32%), 7-10 years (11%), 11-15 years (19%), or more than 15 years (38%). The sample had almost as many female (49%) faculty members as male (51%) members.

Of 19 men interviewed, 14 (73.7%) identified as White, and of 18 women interviewed, 17 (94.4%) identified as White; these racial/ethnic and gender distributions were much higher than the national average for U.S. colleges and universities. The most common academic fields that faculty taught in were social and behavioral studies (30%) and arts and humanities (27%) while far fewer were teaching mathematics (3%) or interdisciplinary studies (3%).

Coding procedures

Manifest and latent coding methods were used for analytic reduction and a systematic interpretation of underlying patterns in the student focus group data logs. Transcripts of the sessions, recorded using Otter.ai software, were also used once participants' names were removed from the files and machine transcription errors were fixed. Focus groups, not individual student responses, were used as the unit for our coding. Eight coding properties were used to analyze comments in all 16 focus groups. These properties were intended to capture what kinds of experiences students, in their own words, had with

computer algorithms. In cases where students in a single group mentioned a concern, e.g. "the next generation," more than once in a session, we only counted it once in our final coding results.

Krippendorff's alpha (KALPHA), considered the most rigorous means of testing intercoder reliability, was run on two pilot test round of focus group logs and coded by two PIL researchers. KALPHA takes into account chance agreement among content analysis coders. While there is no universally accepted standard for intercoder reliability using Krippendorff's alpha, communications researchers have suggested that a coefficient between 0.81 and 0.99 is "almost perfect," between 0.61 and 0.80 is "substantial," and 0.41 to 0.60 is "moderate." Two pilot coding rounds of three interview logs each were used. During the second pilot round, the coding practices reached the acceptable reliability level of 0.84. Thereafter, we coded the focus group logs using eight individual properties for "concerns."

Methodological limitations

There are challenges associated with the use of interviews and focus groups in any research study. For instance, the usefulness of qualitative data collected from interview methodologies depends on participants' provision of accurate and complete answers. Accordingly, the interviewer must endeavor to establish trust and rapport with participants.

Bias on both sides of this kind of exchange is always a formidable issue, too. Bias can be readily introduced in the way the interviewer frames a question, or the way in which a respondent interprets and then answers a question. To enhance the reliability of our interview technique and the consistency of the questions we asked, we used a scripted interview with both student and faculty participants. The script was piloted and a few small changes were made to the wording before the focus groups and interviews began.

Another issue is self-selection of volunteer samples like the ones used in this study, since those that choose to participate in a study may have a special interest in a research topic. While taking all of these limitations into account, we point to the main purpose of qualitative research: interviews are not necessarily used to produce generalizable findings about a sample; rather interviews are used to arrive at a deep understanding of a specific situation, as respondents decide to report them. Despite making every attempt to address these limitations, we acknowledge that future research is required to confirm these findings. Therefore, our findings should not be viewed as comprehensive, but rather as part of our ongoing research about the impact of content from algorithm-driven platforms on information quality.

Table 3: Description of Student Focus Group Sample

Major	Count	Percent
Architecture and Engineering	3	2.91%
Arts and Humanities	14	13.59%
Business Administration	10	9.71%
Computer Science	1	0.97%
Education	7	6.80%
Mathematics	3	2.91%
Occupational Training	7	6.80%
Social or Behavioral Sciences	18	17.48%
Life or Physical Sciences	9	8.74%
Undeclared	4	3.88%
Multiple	18	17.48%
Other	9	8.74%
Total	103	100.00%
Class standing		
Sophomore or second-year student	30	29.13%
Junior or third-year student	32	31.07%
Senior or fourth-year student	29	28.16%
Fifth-year student or beyond	12	11.65%
Total	103	100.00%
Age		
18 years old	15	14.56%
19-20 years old	42	40.78%
21-22 years old	27	26.20%
23-25 years old	9	8.74%
Over 25 years old	9	8.74%
Prefer not to answer	1	0.97%
Total	103	100.00%
Gender		
Female	69	66.99%
Male	31	30.10%
Prefer not to answer	1	0.97%
Total	103	100.00%
Race/Ethnicity		
Asian	6	5.83%
Black/African American	5	4.85%
Hispanic/Latino	4	3.88%
Pacific Islander	2	1.94%
White/Caucasian	74	71.84%
Multiple (more than one selected)	11	10.68%
Prefer not to answer	1	0.97%
Total	103	100.00%

Table 4: Description of Faculty Interview Sample

Disciplinary expertise	Count	Percent
Arts & Humanities	10	27.03%
Business Administration	2	5.41%
Computer Science	2	5.41%
Education	3	8.11%
General Education	1	2.70%
Interdisciplinary	1	2.70%
Life or Physical Sciences	6	16.22%
Mathematics	1	2.70%
Social or Behavioral Sciences	11	29.73%
Total	37	100.00%
Years at institution		
1-6 years	12	32.43%
7-10 years	4	10.81%
11-15 years	7	18.92%
More than 15 years	14	37.84%
Total	37	100.00%
Age		
30-35 years old	5	13.51%
36-41 years old	3	8.11%
41-45 years old	5	13.51%
45-50 years old	1	2.70%
50-55 years old	7	18.92%
Over 55 years old	15	40.54%
Prefer not to answer	1	2.70%
Total	37	100.00%
Gender		
Female	18	48.65%
Male	19	51.35%
Total	37	100.00%
Race/Ethnicity		
White/Caucasian	31	83.78%
Asian	2	5.41%
Multiple (more than one selected)	3	8.11%
Prefer not to answer	1	2.70%
Total	103	100.00%

Keywords and Definitions

Algorithm — the set of logical rules used to organize and act on a body of data to solve a problem or to accomplish a goal that is usually carried out by a machine. An algorithm is typically modeled, trained on a body of data, and then adjusted as the results are examined. Because algorithms are generally processed by computers and follow logical instructions, people often think of them as neutral or value-free, but the decisions made by humans as they design and tweak an algorithm and the data on which an algorithm is trained can introduce human biases that can be compounded at scale. Humans who interact with an algorithm may also find ways to influence the outcomes, as when a marketer finds ways to push a website up in the results of a search through search engine optimization (SEO).

Algorithmic justice — the application of principles of social justice and applied ethics to the design, deployment, regulation, and ongoing use of algorithmic systems so that the potential for harm is reduced. Algorithmic justice promotes awareness and sensitivity among coders and the general public about how data collection practices, machine learning, AI, and algorithms may encode and exacerbate inequality and discrimination.

Algorithmic literacy — a subset of information literacy, algorithmic literacy is a critical awareness of what algorithms are, how they interact with human behavioral data in information systems, and an understanding of the social and ethical issues related to their use.

Artificial intelligence (AI) — a branch of computer science that develops ways for computers to simulate human-like intelligent behavior, able to interpret and absorb new information for improved problem-solving, and recognize patterns. Examples include training robots, speech recognition, facial recognition, and identifying objects such as traffic signs, trees,

and human beings necessary for self-driving cars. AI relies on machine learning capabilities and training data. Humans are involved in creating or collecting sets of training data (e.g., employing low-wage workers abroad to identify objects on computer screens to provide data for autonomous vehicle navigation). Bias may be built into machine learning (e.g., by using criminal justice data sets for risk assessment in predictive policing). Machines can be trained to learn from experience but common sense and recognizing context are difficult, thus limiting the ability of computer programs to perform tasks such as distinguishing hate speech from colloquial humor or sarcasm.

Attention economy — since our attention is a limited resource and every person only has so much of it, companies (both platforms and people who use the platforms to sell, entertain, or persuade) try to engage and keep people's attention. This rewards clickbait and influences the design of algorithms and platforms to maximize time spent online.

Big data — a set of technological capabilities developed in recent years which, when used in combination, allows for the continuous gathering and processing of large volumes of fine-grained and exhaustive data drawn from multiple sources to be combined and analyzed continuously.

Data exhaust — information incidentally generated as people use computers, carry cell phones, or have their behavior captured through surveillance which becomes valuable when acquired, combined, and analyzed in great detail at high velocity.

Machine learning — the use of algorithms, data sets, and statistical modeling to build models that can recognize patterns to make predictions and interpret new data. The purpose of machine

learning is to enable computers to automate analytical model-building so computers can learn from data with little human intervention.

Personalization — the process of displaying search results or modifying the behavior of an online platform to match an individual’s expressed or presumed preferences, established through creating digital profiles and using that data to predict whether and how an individual will act on algorithmically selected information. This process drives targeted digital advertising and has been blamed for exacerbating information silos, contributing to political polarization and the flow of disinformation. Ironically, to consider information “personal” implies it is private, but personalization systematically strips its targets of privacy.

Platform — an ambiguous term that means both software used on personal computers and software deployed online to provide a service, such as web search, video sharing, shopping, or social interaction. Often these systems use proprietary algorithms to mediate the flow of information while enabling third parties to develop apps, advertising, and content, thus becoming digital spaces for the individual performance of identity online, data-driven persuasion (commercial as well as political), and group formation through social interaction. In this report, we use the term to refer to “internet giants” such as Google, YouTube, Instagram, and Facebook and others mentioned by students in our focus group sessions.

About PIL

Project Information Literacy (PIL) is a nonprofit research institute in the San Francisco Bay Area that studies what it is like to be a student in the digital age. In a series of 11 groundbreaking scholarly research studies, PIL has investigated how U.S. college students and recent graduates utilize research skills, information competencies, and strategies for completing course work, engaging with news, and solving information problems in their everyday lives and the workplace. Research findings and recommendations from PIL studies have informed and influenced the thinking and practices of diverse constituencies from all over the world from those in higher education, public libraries, newspapers, and the workplace.

projectinfolit.org
@projectinfolit

About the algorithm study

Preferred citation format: Alison J. Head, Barbara Fister, and Margy MacMillan, *Information literacy in the age of algorithms: Student experiences with news and information, and the need for change* (15 January 2020), Project Information Research Institute, <https://projectinfolit.org/publications/algorithm-study/>

Abstract: This report presents findings about how college students conceptualize the ever-changing online information landscape, and navigate volatile and popular platforms that increasingly employ algorithms to shape and filter content. Researchers conducted 16 focus groups with 103 undergraduates and interviews with 37 faculty members to collect qualitative data from eight U.S. colleges and universities from across the country. Findings suggest that a majority of students know that popular websites, such as Google, YouTube, Instagram, and Facebook, use algorithms to

collect massive amounts of their personal data, but still find sites too useful to abandon. Many are indignant about websites that mine their clicks to sell them products, but resigned to the powers of an unregulated media environment. Some students, however, used practical strategies to protect their privacy and “confuse algorithms,” learned more often from peers than in classes. An abundance of choice for online information left many skeptical and distrustful of news encountered on algorithm-driven platforms. While some students worried about the “creepy” way ads follow them around the internet, others were concerned that automated decision-making systems reinforce societal inequalities. Discussions with students and faculty indicated that understanding and managing the torrent of information flowing through search engines and social media is rarely mentioned in the classroom, even in courses emphasizing critical thinking and information literacy. A critical review of a decade of research from Project Information Literacy (PIL) about how students conduct course and everyday life research, and what that means for educators and librarians, provides context to these new findings. Four recommendations are provided for educators, librarians, administrators, and journalists working to promote truth and prepare students for a changing and challenged world.

The Algorithm Study Research Report has a Creative Commons (CC) license of “CC BY-NC-SA 4.0.” This license allows others to share, copy, adapt, and build upon the survey data non-commercially, as long as the source — Project Information Literacy — is credited and users license their new creations under the identical terms.



Supplementary resources

There is a landing page with additional resources from the PIL algorithm study. All of these materials are open access and can be used without permission from PIL, https://www.projectinfolit.org/algo_study.html

Funding support

The John S. and James L. Knight Foundation has generously funded this research with support from Harvard Graduate School of Education (HGSE), Electronic Resources & Libraries, a leading library conference, and the School of Library and Information Science at the University of South Carolina.

Individual donor support

Suzanne Bernsten
 Anne Deutsch
 Emily Drabinski
 Robin Ewing
 Sarah Fancher
 Alex Hodges
 Trudi Jacobson
 Gayle Jenkins
 David Lankes
 Clarence Maybee
 Panagiotis Metaxas
 Maryanne Mills
 Peter Morville
 Bonnie Tijerina
 Sandy Tijerina
 Sarah Vital
 Scott Walter

Algorithm study research team**Principal investigator**

Alison J. Head, Ph.D.
*Project Information Literacy Executive Director
 Visiting Scholar,
 Harvard Graduate School of Education
 Senior Researcher, metaLAB (at) Harvard*
alison@projectinfolit.org

Co-investigators

Barbara Fister
*Inaugural Scholar-in-Residence, PIL
 Emerita Professor,
 Gustavus Adolphus College, Minnesota*

Margy MacMillan
*Senior Researcher, PIL
 Emerita Professor, Mount Royal University, Canada*

Research analysts

Alaina Bull
*First Year Experience Librarian, The University of
 Washington Tacoma*

Erica DeFrain, Ph.D.
*Assistant Professor, Social Sciences Librarian,
 University of Nebraska-Lincoln Libraries*

Jessica Yurkofsky
*PIL Fellow
 metaLab [at] Harvard*

Editing

Jena Gaines
 Meg McConahey

Graphic design and data analytics

Steven Braun

*Visiting Assistant Professor in the College of Arts,
Media and Design, Northeastern University*

Erica DeFrain, Ph.D.

*Assistant Professor, Social Sciences Librarian,
University of Nebraska-Lincoln Libraries*

Kirsten Hostetler

*Outreach Faculty Librarian,
Central Oregon Community College*

Jessica Yurkofsky

*PIL Fellow
metaLab [at] Harvard***Institutional sample support**

Anne-Marie Deitering,

Oregon State University (OR)

Sarah Fancher,

*Ozarks Technical Community College (MS)*Andrew Haggarty, *Clark University (MA)*Phil Jones, *Grinnell College (IA)*Terri Summey, *Emporia State University (KS)*Garrett Trott, *Corban University (OR)*Celia Rabinowitz, *Keene State College (NH)*Michele Van Hoeck, *California State University**Maritime Academy (CA)***Image attributions****Cover photo**<https://unsplash.com/collections/8467201/people>Photo by [Jan Kolar \(www.kolar.io\)](http://www.kolar.io) on [Unsplash](https://unsplash.com/)**What exactly is information literacy?**<https://unsplash.com/photos/iYcuJQaVTvg>by [Etienne Boulanger](https://unsplash.com/) on [Unsplash](https://unsplash.com/)**When algorithms “get creepy”**US Department of Defense - <https://www.jbsa.mil/News/Photos/igphoto/2001937728/>,Public Domain, <https://commons.wikimedia.org/w/index.php?curid=81108948>**The mainstreaming of extremism**Photo by [Andrea Bellucci](https://unsplash.com/) on [Unsplash](https://unsplash.com/)<https://unsplash.com/photos/mqpwdNwa0o8>**The “pivot generation”**

Photo by Kheng Ho Toh

Acknowledgements

This study would have never been possible without the students, faculty, and librarians that participated in our research activities and generously gave of their time. We also owe a world of thanks to the incredibly dedicated and skilled PIL Research Team: Alaina Bull (The University of Washington Tacoma), Erica DeFrain (University of Nebraska-Lincoln), and Jessica Yurkofsky (metaLAB (at) Harvard). Throughout the “Algo Study” there were many PIL supporters-at-large lending encouragement, help, and suggestions as this project unfolded: Alex Hodges (Harvard Graduate School of Education and Gutman Library), Nate Hill (Metropolitan New York Library Council - METRO), Kirsten Hostetler (Central Oregon Community College), Kei Kawashima-Ginsberg (Tufts University), Takis Metaxas (Wellesley College), Momin Malik (Berkman-Klein Center for Internet and Society), Karen Schneider and Rita Premo (Sonoma State University), Ronald Robertson (Northeastern University), Marcie Rothman, Mandy Shannon (Wright State University), Michele Van Hoeck (California State University Maritime Academy), and John Wihbey (Northeastern University). Also, we are grateful to (and amazed by) the 81 PIL research liaisons at U.S. colleges and universities, who signed up for consideration to fill one of the eight openings open in our institutional sample last spring.

-Alison J. Head, Barbara Fister, and Margy MacMillan
January 15, 2020