FSC Washington Testimony.

"The Future of Identity in Financial Services: Threats, Challenges, and Opportunities,"

The Task Force on Artificial Intelligence
House Financial Services Committee
United States Congress

Anne L. Washington, PhD
Steinhardt School
New York University

September 12 2019

Thank you for inviting me to speak today. My name is Anne Washington. I am an

Assistant Professor of Data Policy in the Department of Applied Statistics[1] at New York

University. Before my career in Academia, I spent eight years in financial services with a data-

driven company in San Francisco. I also spent a decade in the legislative branch working with

many of the data structures and tools used to make this hearing possible. I would also like to

acknowledge that I serve on the Academic Advisory Board of the Electronic Privacy Information

Center, EPIC. I hold an undergraduate degree in computer science from Brown University, a

graduate degree in Library and Information Science from Rutgers University, and a doctorate

from The George Washington University School of Business.

My testimony, today, represents my own views as a public interest technologist. As a

computer scientist and organizational scholar with expertise in open government data, I am part

---

[1] I am in the Steinhardt School of Culture, Education, and Human Development at NYU and the Department of Applied

of a growing movement of people[2] using STEM[3] skills in non-profits and the public sector. My academic specialty is understanding the organizational dynamics that shape the production and consumption of information, especially in organizations that have a public mission.

The courses I teach to graduate students at New York University are the "Management and Ethics of Data" and the "Ethics of Data Science". In my testimony today, I will give you a crash course on data ethics, squeezing two semester-long courses into a five-minute briefing.

Artificial intelligence is not infallible. Even the most successful artificial intelligence systems used by online financial platforms require human input. For Americans to participate equally in our financial system, we need inclusive innovation that is aware of difference. Ignoring AI exceptions in financial services risks excluding many in our society because they are outliers from expectations. Organizations must begin to think about how they will handle future disputes over AI errors.

Artificial intelligence in the financial sector is an ethical, mathematical, and policy issue. To illustrate this, I will elaborate on three main points:

1. Artificial intelligence produces errors. When operating "at scale" even low error rates can impact millions. Errors in financial services will be consequential to specific individuals.

2. Because organizations are more likely to believe their technology systems over the experiences of individuals, individuals need procedures for recourse in the event of processing error.

3. Systems built to consider a broader range of populations must be more fault tolerant of cultural difference to be robust.

---

[2]   Such as Desmond Patton trained in computer science and social work at Columbia University. Dierdre Mulligan trained in Law and teaching in the Berkeley Information School. See Bruce Schneier's Public Interest Technology list.
[3]   Science Technology Engineering and Math

Ethics

The study of ethics concerns itself with questions of appropriate behavior and actions. For centuries, the assumption behind ethics has been that we, as human beings, were driving our actions. Today, we are confronted with computer systems acting on behalf of humans. Ethical questions arise when actions violate the public trust.

Artificial intelligence is a technology that gives organizations an incredible power over individuals. M. Lynne Markus[4] (2016) reminds us that the information on millions of people is in the hands of only a few and those organizations have a "corporate *data* responsibility".[5]

Data technology, such as artificial intelligence, drives all sectors of industry including financial services. Digital material from sensors, transactions, cell phones, networks, social media, and other digital traces feed into systems that generate artificial intelligence. Digital traces like these when reused in new contexts might trigger ethical concerns if not traceable and joined appropriately.

These pipelines into the "data supply chain"[6] are mostly owned and operated by corporate bodies and not individuals. Christine Borgman (2015) argues that digital systems generate not just big data, but also small data, or even no data[7] . These natural inconsistencies can create havoc when data technologists attempt to connect data from different sources. Many of these

---

[4] Markus, M. L. (2016). Obstacles on the Road to Corporate Data Responsibility. In C. R. Sugimoto, H. R. Ekbia, & M. Mattioli (Eds.), Big data is not a monolith (p. 143). MIT Press.
[5] The term Corporate Data Responsibility is based on the well known concept in management of corporate social responsibility. Some scholars are now understanding data is a part of that responsibility within supply chains. See Amaeshi, K. M., Osuji, O. K., & Nnodim, P. (2007). Corporate Social Responsibility in Supply Chains of Global Brands: A Boundaryless Responsibility? Clarifications, Exceptions and Implications. Journal of Business Ethics, 81(1), 223–234. doi: 10.1007/s10551-007-9490-5
[6] Washington, Anne L. (2014). Data Supply Chains. Invited Workshop Leader. The Social, Cultural, & Ethical Dimensions of "Big Data" sponsored by the White House Office of Science and Technology Policy (OSTP). Afternoon Breakout Sessions. Hosted by Data & Society Research Institute and the New York University (NYU) Information Law Institute. March 17. New York, NY. http://www.whitehouse.gov/issues/technology/big-data-review
[7] Borgman, C. L. (2015). Big data, little data, no data: Scholarship in the networked world. http://lccn.loc.gov/2014017233.

technologies are perfectly legal and necessary for innovative growth, however ethical questions remain. Those in power who use this data must be reminded that data can make people vulnerable.

Artificial intelligence (AI) has grown our economy by driving economies of scale. Its efficiency provides gains in productivity and precision (Dhar, 2013; Halevey & Norvig, 2009). AI, however, can also obscure policies (Eubanks, 2018), and exacerbate bureaucracy (Peeters & Schuilenburg, 2018) amongst other concerns (Rossi, 2019; Wiggen, 2017). The tension between pragmatic efficiency and the moral tug of appropriate action plagues adoption of AI technologies by governments.

The artificial intelligence[8] I discuss here, today, is data technology that enables oversight, automates decisions, or augments observations over large streams of data. Data technology includes data science, machine learning, predictive analytics, evidence-based policy, and computational tools based on the consumption and analysis of large quantities of information. Usually, these systems work with algorithms that sort, rank, search, and calculate in order to generate consistent outcomes. On the surface, these systems appear to be neutral, mechanical, and routine-driven, but when placed within human societies they can have substantial repercussions within our daily lives. The computer scientist Meredith Broussard says that these socially agnostic systems are not robust enough and labels them as "artificial unintelligence".[9]

The power of data technology is derived from the amount of data it uses. When data technology struggles to identify individuals in a database, the solution is often to combine more

---

[8]   Intelligence can take many forms and has changed definitions over the years. Belkin, N. J. (1996). Intelligent information retrieval: Whose intelligence? In ISI - International Symposium for Information Science: Vol. 96. Proceedings of the Fifth International Symposium for Information Science (pp. 25–31). and Gardner, H. (1983). The Idea of Multiple Intelligences. In Frames of Mind: The theory of multiple intelligences (pp. 3–11). New York: Basic Books Inc.
[9]   Broussard, Meredith 2018. Artificial Unintelligence: How Computers Misunderstand the World. MIT Press, Boston, Mass.

databases into decision making. Amassing data in this way makes individuals entirely too visible (Rocher, 2019; Sweeney, 2013). Privacy and the "politics of real names,"[10] danah boyd tells us, are real concerns. At the core of these concerns are questions not only of social categorization[11] (Cherng, 2017) but also of technical abstraction (Walsh, 1992).

Ethics programs, like the one at NYU, want to help build better data systems. By baking privacy, security, and usability into the design of our AI systems, we can build a more responsible and ethical data environment like the solutions proposed by (Shilton, 2013) and (Cranor & Garfinkel, 2005)[12]. Others, such as the scholars at the Ostrom Center for Data Commons, are using the work of Nobel-prize winning economist, Elinor Ostrom, to better understand the ethics of knowledge commons (Raymond, 2018).

Data Ethics In Real Life

Every student of data ethics understands that large populations, coincidence, and cases of mistaken identity can confound the "trustworthiness" of AI systems.

Large Data Sets

Current AI-based systems, including financial systems incorporating AI, are not ready to

---

[10] boyd, danah. (2010, October). Why Privacy Is Not Dead. Technology Review, 113(5), 10–11.
boyd, danah. (2012). The politics of "real names." Communications of the ACM, 55(8), 29–31. doi: 10.1145/2240236.2240247
[11] The classic text on this is Bowker, G. C., & Star, S. L. (1999). Sorting things out: Classification and its consequences. MIT PRESS. See also   Bowker, G. C., & Star, S. L. (2000). Invisible Mediators of Action: Classification and the Ubiquity of Standards. Mind, Culture, and Activity, 7(1–2), 147–163. doi: 10.1080/10749039.2000.9677652
Suchman, L. (1993). Do categories have politics? Computer Supported Cooperative Work (CSCW), 2(3), 177–190. doi: 10.1007/BF00749015
[12] See Schlesinger, A., O'Hara, K. P., & Taylor, A. S. (2018). Let's Talk About Race: Identity, Chatbots, and AI. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, 1–14. https://doi.org/10.1145/3173574.3173889 and Levy, K & Barocas, S ."DESIGNING AGAINST DISCRIMINATION" https://scholarship.law.berkeley.edu/btlj/vol32/iss3/5/ doi.org/10.15779/Z38BV79V7K

disambiguate enormous sets of people. The USA current has close to 330 million people. Some technology platforms have more users than the populations of countries. These large Internet platforms have specific authentication methods to identify a user who logs in that includes active involvement of the individual logging in. Financial service data, which travels between multiple institutions, is harder to track down and does not include the end-user actively as part of authentication and identification. This leads to loosely coupled systems that introduce noise into financial profiles. Evelyn Ruppert (2009, 2011) calls the digital traces that represent us a "data double". Our data double is similar to us but not exactly. Daniel Solove (2004) calls these data traces an "unauthorized biography" that contains some true things but lots of noise and innudedo that is not true. The scale of users and their data can introduce error into the decisions made by financial sector AI.

The Birthday Problem

Coincidences[13] are not a surprise to any student of statistics. Basic math theory tells us that what we expect is rare may be more likely than we think. The inquiry known as the Birthday Problem[14] asks: how many people are needed in a room for a good chance that two people are born on the same day of the year? Surprisingly the number is just 23. There is a 50% chance that in a room of 23 people, given true randomness, two people have the same birthday. In a room of 75 people, the chances are over 99% and a 1/3 chance that three people do. When two items resolve to the same set of information, computer scientists building a hash algorithm know this as a "hash collision".

The Birthday Problem has real world implications when we use this information to

---

[13] Stewart, I. (1998). What a Coincidence! - MATHEMATICAL RECREATIONS. Scientific American, 2. doi: DOI: 10.1038/scientificamerican0698-95
[14] Weisstein, E. W. (2019, September 6). Birthday Problem. Retrieved September 1, 2019, from MathWorld Wolfram website: http://mathworld.wolfram.com/BirthdayProblem.html

disambiguate identities. A classic example is the problem of watch lists that permit entry or deny

services. Jeff Jonas, a pioneer in entity recognition, has explored this tension between privacy

and recognition in a famous paper about the terrorist watch list. The authors conclude that

actionable information is more important than aggregate lists that violate civil liberties (Harper

& Jonas, 2006). The legal scholar Margaret Hu (2015) goes into extensive detail about these

lists and their impact on people's lives.

These problems are not new (Solove, 2001; 2004) nor unknown to computer scientists and

statisticians (Becker, 2006). What is new, however, is that these materials are moving from

identification into action in ways that can aggregate a single mistake into an ongoing situation.

The data supply chain moves not in one direction but in circles exacerbating mistakes.

Examples of mistaken identity

People have a difficult time fighting these lists once their names are on them.

- Jennifer Norris[15] of Boston was in danger of losing her job because of the inability to

    resolve a dispute about her identity. Her work required a driver's license and only after

    consulting her Congressman, Capuano of Massachusetts' 7th district and a local news

    agency was the problem resolved.

- Kathleen Casey[16], a pharmacy technician, lost her apartment in 2011 when a system

    confused her with someone else. It is important to note that some industries[17] such as

    retail pharmacy stores used informal lists to exclude any job candidate accused of theft.

---

[15] Kath, R. (2018, Apr 11) I-Team: Mistaken Identity Causes Bureaucratic Nightmares For Drivers. CBS News Boston . https://boston.cbslocal.com/2018/04/11/wbz-tv-i-team-drivers-bureaucratic-nightmares-mistaken-identity-federal- database/
[16] Liedtke, Michael (2011, Dec 16)   How A Clerical Error Put A Woman On The Streets. Business Insider / AP   . . https://www.businessinsider.com/mistaken-identity-put-this-woman-on-the-streets-2011-12
[17] Clifford, S., & Silver-Greenberg, J. (2013). Retailers Use Databases to Track Worker Thefts. The New York Times. And Knaub, Kelly (2014, Aug 4) "LexisNexis, Retail Workers Get Nod For $2.38M Settlement," Law 360. http://www.law360.com/articles/563583/lexisnexis-retail-workers-get-nod-for-2-38m-settlement.

- A teacher in Maryland[18] could not pursue her chosen livelihood because bad data continually haunted her in a job that requires continuous recertification.

The astounding case of Lisa S. Davis[19], the novelist, who wrote about her experience of encountering her data double in official documents for 18 years and finally meeting her. For years, their addresses were confused and they would get mail for each other. They had the same day of birth, the same year of birth, and not only the same middle initial, but the same middle name. Most systems have a hard time if not impossible time disambiguating them. They assume it is one person who perhaps has just moved to a new address. This data double story has more resonance in this case because the two women are different colors and live in neighborhoods with different policing behaviors. They are both in New York State so their information has a higher chance to be co-mingled in databases.

Resolving disputes

Davis (2017) relates her story of having information that would show that her experience and paper traces verified who she was. Her lived experience was no match to the certainty of a computer. She was assumed to be a liar and told to plead guilty to pay and clear the traffic violations.

Organizations tend to trust their computer systems over the customers' experience. Individuals with a wrong match, who are outliers, who clearly can identify a flaw in the system, are perceived as liars. Humans take the blame after a systems provides an answer. People with

---

[18]   Meyer, Eugene L. (1997, Dec 15) "Md. Woman Caught in Wrong Net; Data Errors Link Her to Probes, Cost 3 Jobs," Washington Post C1. As discussed in Solove, D. (2004). The digital person: Technology and privacy in the Information Age. New York: New York University Press.
[19]   Davis, Lisa Selin (2017, Apr 3) For 18 years, I thought she was stealing my identity. Until I found her . The UK Guardian https:// www.theguardian.com/us-news/2017/apr/03/identity-theft-racial-justice

lived experience that contradicts the artificial intelligence face significant challenges. It is like watching a toy robot go towards the corner and march in place endlessly.

Technologists building these systems want to learn this feedback. Businesses do not have a financial incentive to incrementally fix small errors. Any policy or best practice would give technologists inside organizations the leverage they need to spend their time fixing the errors. This feedback once incorporated could help prevent similar mistakes from being repeated later. This agile approach with feedback would help to incrementally improve the technology.

Individuals should have recourse in these situations. It is mathematically certain that collisions will occur. Without any form of redress, innovation will stall and many people will be locked out of financial systems.

It is important to note that these stories all focus on individuals but one-person Internet shops[20] that rely on technology infrastructure are even more vulnerable. Owner-operator and new entrepreneurs, and small business who are establishing their validity in markets have high risks if locked out of financial capital.

POLICY

My remarks on policy alternatives will be the most brief. Legal scholarship in this area is extensive especially in data[21] used in policing and court data[22] system. The scholarship of

---

[20]   BBC News (2019, Aug 15) 'My Instagram got hacked and I lost my business' - BBC News . . .
https://www.bbc.com/news/business-49397038
[21]   See works by Julie Cohen and Paul Ohm. Especially: Cohen, Julie E. (2012). What privacy is for. Harvard Law Review, 126, 1904. ; Ohm, Paul. (2009). Broken promises of privacy: Responding to the surprising failure of anonymization. UCLA l. Rev., 57, 1701. and Ohm, Paul. (2011). The fourth amendment in a world without privacy. Miss. LJ, 81, 1309.
[22]   See Wexler, R. (2017, June 13). When a Computer Program Keeps You in Jail. The New York Times, p. A27. Wexler, R.

Danielle K. Citron and Frank A. Pasquale [23] have covered many plausible solutions.

I suspect that there will be a debate over the feasibility of establishing recourse. As usual, some will suggest that new regulations should be put in place. It would be logical to extend the Fair Credit Reporting Act (FCRA) into the 21st century and acknowledge the role of data sources. Others will suggest that self-monitoring would be sufficient. It makes sense to allow innovation to develop without unnecessary constraints since the future of these technologies is hard to foresee. Associations[24] and industry cooperatives[25] could continue to establish best practices across the field.

In my opinion, neither of these traditional responses gets to the heart of the issue which is that data-driven organizations need to establish internal data policy that matches their values with the business model. Data-driven organizations might run on a variety of business models (Shapiro & Varian, 1998) so it is a management decision what policy best matches those goals. I see these concerns as extensions of early conversations about organizational memory (Ackerman, 2000; Anand, 1998) that asked what information[26] should organizations keep as technology became ubiquitous.

(2018). Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. Stanford Law Review. Selbst, A. D. (2017). Disparate Impact in Big Data Policing. Ga. L. Rev., 52, 109.; And Hu, M. (2017). Algorithmic Jim Crow. Fordham Law Review, 86(2), 633-696. And Eaglin, J. M. (2017). Constructing Recidivism Risk. Emory Law Journal, 67(1), 59–122.

[23]   See Pasquale, F. (2015a). Reforming the Law of Reputation. Loyola University Chicago Law Journal, 47, 25. Pasquale, F. (2015b). The black box society: The secret algorithms that control money and information. Pasquale, F. (2017). Toward a Fourth Law of Robotics: Preserving Attribution, Responsibility, and Explainability in an Algorithmic Society. Ohio State Law Journal, 78(5), 1243–1255. Pasquale, F. A. (2018). A Rule of Persons, Not Machines: The Limits of Legal Automation. George Washington Law Review. Pasquale, F. A., & Citron, D. K. (2014). Promoting Innovation While Preventing Discrimination: Policy Goals for the Scored Society. Washington Law Review, 89, 1413–1424.
 And Citron, D. K. (2008a). Cyber Civil Rights. Boston University Law Review, 89, 61–125.
Citron, D. K. (2008b). Technological Due Process. Washington University Law Review, 85(6), 1249–1313.
Citron, D. K. (n.d.). Fulfilling Government 2.0's Promise with Robust Privacy Protections. The George Washington Law Review, 78, 24.

[24]   A group like the Public Policy council for the Association of Computing Machinery is an ideal place to work out many of these policy issues with experienced technologists.

[25]   An industry membership group like , such the Partnership on AI, the Future of Privacy Forum, or the Data Coalition, can share invaluable and realistic advice when confronting these issues in themidsts of operations.

[26]   Agar, J. (2006). What Difference Did Computers Make? Social Studies of Science, 36(6), 869–907.

One critical role for public policy is data standards. Governments could establish data standards that would relieve the burdens of anyone trying to investigate issues across firms. Standardized data structures could also a mechanism to trigger retrospective tracking for regulators, e-discovery, or internal business intelligence. The reuse and exchange of digital material is complicated by many social and organizational challenges (Borgman, 2000; Bowker, 1996; Edwards, 2011; Fedorowicz, 2010; Markus, 2006). A solid internal information policy (McClure 1989; Robinson, Yu, Zeller, & Felton, 2008) is critical for any data-driven organization. For example, in the public sector the 2014 DATA Act produced a stable data infrastructure across all agencies that made later analysis, correction, and innovation possible. Digital government scholars such as Sharon Dawes (1996, 2010), Theresa A. Pardo (2012), Marijn Janssen (2016), Lemuria Carter (2018), Paul Jaeger & John Bertot (2010) have written extensively about the importance of data structures in government transparency.

Governments often neglect that their greatest power in public policy is mandating data infrastructure. Identity standards for financial services would greatly serve to expedite the adoption of artificial intelligence that benefits wide audiences.

Summary

Artificial intelligence, often implemented to save labor costs, will still require human labor to handle anticipated exceptions. A dispute resolution process solves two problems: procedural justice and technology improvement. First, it establishes a procedure to preserve the sanctity of human experience in situations where organizations may be more likely to trust the AI over a customer. Second, it provides the necessary feedback for incremental improvement of the technology.

Artificial intelligence will have its exceptions and people need procedures to assert the

authority of their lived experience over the authority of the numbers.

BIBLIOGRAPHY

Ackerman, M. S., & Halverson, C. A. (2000). Reexamining organizational memory. Communications of the ACM, 43(1), 58–64. doi: 10.1145/323830.323845

Aitkin, R. (2017). 'All data is credit data': Constituting the unbanked. Competition & Change, 21(4), 274–300. doi: 10.1177/1024529417712830

Azerrad, David (2019) Back Row America. https://www.heritage.org/poverty-and-inequality/commentary/back-row-america

Becker, J. P., & Wallace, J. (2006). Eighty Ways to Spell Refrigerator. Proceedings of the Thirty-First Annual SAS Users Group International Conference. Retrieved from http://www2.sas.com/proceedings/sugi31/074-31.pdf

Belkin, N. J. (1996). Intelligent information retrieval: Whose intelligence? In ISI - International Symposium for Information Science: Vol. 96. Proceedings of the Fifth International Symposium for Information Science (pp. 25–31). Retrieved from http://comminfo.rutgers.edu/tipster3/iirs.html

Benjamin, R. (2019). Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life. Durham, NC: Duke University Press.

Benjamin, R. (2019). Race After Technology: Abolitionist Tools for the New Jim Code. Cambridge, UK: Polity.

Berlins, M. (2008, November 2). A Kafkaesque excuse for ignorance of the law. The Guardian. Retrieved from http://www.theguardian.com/uk/2008/nov/03/law-kafka-transparency-marcel-berlins/print

Borgman, C. L. (2000). From Gutenberg to the global information infrastructure: Access to information in the networked world. Cambridge, MA: MIT Press.

Borgman, C. L. (2015). Big data, little data, no data: Scholarship in the networked world. .

Bowker, G. C. (1996). The history of information infrastructures: The case of the international classification of diseases. Information Processing & Management, 32(1), 49–61. doi: 10.1016/0306-4573(95)00049-M

Bowker, G. C., & Star, S. L. (1999). Sorting things out: Classification and its consequences.

Bowker, G. C., & Star, S. L. (2000). Invisible Mediators of Action: Classification and the Ubiquity of Standards. Mind, Culture, and Activity, 7(1–2), 147–163. doi: 10.1080/10749039.2000.9677652

boyd, danah. (2010, October). Why Privacy Is Not Dead. Technology Review, 113(5), 10–11.

boyd, danah. (2012). The politics of "real names." Communications of the ACM, 55(8), 29–31. doi: 10.1145/2240236.2240247

Broussard, Meredith 2018. Artificial Unintelligence: How Computers Misunderstand the World. MIT Press, Boston, Mass.

Bryson, J. J., Diamantis, M. E., & Grant, T. D. (2017). Of, for, and by the people: The legal lacuna of synthetic persons. Artificial Intelligence and Law, 25(3), 273–291. doi: 10.1007/s10506-017-9214-9

Cadwalladr, C. (2018, December 23). Our Cambridge Analytica scoop shocked the world. But the whole truth remains elusive. The Guardian.

Chamorro-Premuzic, T. (2015, September 24). How different are your online and offline personalities? The Guardian. Retrieved from https://www.theguardian.com/media-network/2015/sep/24/online-offline-personality-digital-identity

Cherng, H.-Y. S. (2017). The Color of LGB: Racial and Ethnic Variations in Conceptualizations of Sexual Minority Status. Population Review, 56(1). doi: 10.1353/prv.2017.0002

Chouldechova, A., & G'Sell, M. (2017, June 30). Fairer and more accurate, but for whom? Presented at the 2017 Workshop on Fairness, Accountability, and Transparency in Machine Learning (FAT/ML 2017), Halifax, Canada.

Chupapados Data https://chupadados.codingrights.org/en/they-are-stalking-you-to-calculate-your-score/

Citron, D. K. (2008a). Cyber Civil Rights. Boston University Law Review, 89, 61–125.

Citron, D. K. (2008b). Technological Due Process. Washington University Law Review, 85(6), 1249–1313.

Clifford, S., & Silver-Greenberg, J. (2013). Retailers Use Databases to Track Worker Thefts. The New York Times.

Cohen, Julie E. (2012). What privacy is for. Harvard Law Review, 126, 1904.

Cramer, Katherine J. (2017, Jun 19). The Great American Fallout: How Small Towns Came to Resent Cities. The Guardian.

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. Organization Science, 10(1), 104–115. doi: 10.1287/orsc.10.1.104

Culnan, M. J., & Williams, C. C. (2009). How Ethics Can Enhance Organizational Privacy: Lessons from the Choicepoint and TJX Data Breaches. MIS Quarterly, 33(4), 673–687.

Davis, Lisa Selin (2017, Apr 3) For 18 years, I thought she was stealing my identity. Until I found her . The UK Guardian https:// www.theguardian.com/us-news/2017/apr/03/identity-theft-racial-justice

Dawes, S. S. (1996). Interagency Information Sharing: Expected Benefits, Manageable Risks. Journal of Policy Analysis and Management, 15(3), 377–394.

Dawes, S. S. (2010). Stewardship and usefulness: Policy principles for information-based transparency. Government Information Quarterly, 27(4), 377–383. doi: 10.1016/j.giq.2010.07.001

Desmond, Ruth. (2010). Consumer Credit Reports and Privacy in the Employment Context: The Fair Credit Reporting Act and the Equal Employment for All Act. University of San Francisco Law Review, 44(4), 6.

Dhar, V. (2013). Data Science and Prediction. Communications of the ACM, 56(12), 64–73. doi: 10.1145/2500499

Eaglin, J. M. (2017). Constructing Recidivism Risk. Emory Law Journal, 67(1), 59–122.

Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science friction: Data, metadata, and collaboration. Social Studies of Science, 41(5), 667–690. doi: 10.1177/0306312711413314

Egan, P. (2017, July 31). Data glitch was apparent factor in false fraud charges against jobless claimants. Retrieved September 7, 2019, from Detroit Free Press website: https://www.freep.com/story/news/local/michigan/2017/07/30/fraud-charges-unemployment-jobless-claimants/516332001/

Eubanks, V. (2018, January 15). A Child Abuse Prediction Model Fails Poor Families. Wired. Retrieved from https://www.wired.com/story/excerpt-from-automating-inequality/

Fedorowicz, J., Gogan, J. L., & Culnan, M. J. (2010). Barriers to Interorganizational Information Sharing in e-Government: A Stakeholder Analysis. The Information Society, 26(5), 315–329. doi: 10.1080/01972243.2010.511556

Fergus, D. and Boyd, T. (2014) Banking without borders. Kalfou, 1(2), 7-28. doi: 10.15367/kf.v1i2.30

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6), 381.

Floridi, L. (2009). Network Ethics: Information and Business Ethics in a Networked Society. Journal of Business Ethics, 90(Supplement 4), 649–659. doi: doi:10.1007/s10551-010-0598-7

Freelon, D., McIlwain, C., & Clark, M. (2018). Quantifying the power and consequences of social media protest. New Media & Society, 20(3), 990–1011. doi: 10.1177/1461444816676646

Gardner, H. (1983). The Idea of Multiple Intelligences. In Frames of Mind: The theory of multiple intelligences (pp. 3–11). New York: Basic Books Inc.

Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. Science, 122(3159), 108–111. doi: 10.1126/science.122.3159.108

Garriga, E., & Melé, D. (2012). Corporate Social Responsibility Theories: Mapping the Territory. In Citation Classics from the Journal of Business Ethics (pp. 69–96). Dordrecht, NETHERLANDS: Springer.

Gelman, A., & Loken, E. (2014). Ethics and Statistics: The AAA Tranche of Subprime Science. CHANCE, 27(1), 51–56.

Gelman, A., & Palko, M. (2013). Ethics and Statistics: The War on Data. CHANCE, 26(1), 57–60.

Gioia, D. A. (1992). Pinto Fires and Personal Ethics: A Script Analysis of Missed Opportunities. Journal of Business Ethics, 11(5/6), 379–389.

Giunchiglia, F., Villafiorita, A., & Walsh, T. (1997). Theories of Abstraction. Artificial Intelligence, 57.

Gladwell, M. (2008). Outliers: The Story of Success.

Gladwell, M. (2011). The Order of Things. New Yorker, 87(1), 68–75. Retrieved from a9h.

Gorman, M. E., Winkler, P. W., & American Library Association. (1998). Anglo-American cataloguing rules (Vol. 2nd). Ottawa. Chicago: Canadian Library Association ; American Library Association.

Gray, J., & Davies, T. G. (2015, May 27). Fighting Phantom Firms in the UK: From Opening Up Datasets to Reshaping Data Infrastructures? Presented at the Open Data Research Symposium at the 3rd International Open Government Data Conference, Ottawa. Retrieved from http://papers.ssrn.com/abstract=2610937

Griffith, J. C. (2001). Congress' legislative information systems: THOMAS and the LIS. Government Information Quarterly, 18(1), 43–60. doi: 10.1016/S0740-624X(00)00066-6

Hacking, I. (2007). Kinds of People: Moving Targets. In P. J. Marshall (Ed.), Proceedings of the British Academy, Volume 151, 2006 Lectures (pp. 285–318). Oxford University Press / British Academy.

Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. IEEE Intelligent Systems, 24(2), 8–12. doi: 10.1109/MIS.2009.36

Hedenus, A., & Backman, C. (2017). Explaining the Data Double: Confessions and Self-Examinations in Job Recruitments. Surveillance & Society, 15(5), 640–654. doi: 10.24908/ss.v15i5.6380

Herkenhoff, K., Phillips, G., & Cohen-Cole, E. (2016). The impact of consumer credit access on employment, earnings and entrepreneurship (No. w22846). National Bureau of Economic Research.

Hess, C., & Ostrom, E. (2011). Understanding knowledge as a commons: From theory to practice. Retrieved from http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=12504

Ho, K. (2009) Leveraging Dominance and Crises through the Global. In Liquidated: An Ethnography on Wall St. doi: 10.1215/9780822391371-008

Ho, K. (2018) Markets, myths, and misrecognitions: Economic populism in the age of financialization and hyperinequality. Economic Anthropology 5(1), 148-150

Hoofnagle, C. J. (2009). Internalizing Identity Theft. UCLA Journal of Law and Technology, 13(2), 38.

Hu, M. (2017). Algorithmic Jim Crow. Fordham Law Review, 86(2), 633-696.

Jaeger, P. T., & Bertot, J. C. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information. Government Information Quarterly, 27(4), 371–376

Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. Government Information Quarterly, 33(3), 371–377. doi: 10.1016/j.giq.2016.08.011

Jonas, J., and J. Harper. 2006. Effective counterterrorism and the limited role of predictive data mining. Policy Analysis. Washington D.C.: Cato Institute, December 11.

Kath, R. (2018, Apr 11) I-Team: Mistaken Identity Causes Bureaucratic Nightmares For Drivers. CBS News Boston https:// boston.cbslocal.com/2018/04/11/wbz-tv-i-team-drivers-bureaucratic-nightmares-mistaken-identity-federal- database/

Katz, D. M., & Bommarito, M. J. (2014). Measuring the Complexity of the Law: The United States Code. Artificial Intelligence and Law, 22(4), 337–374. doi: 10.1007/s10506-014-9160-8

Knaub, Kelly (2014, Aug 4) "LexisNexis, Retail Workers Get Nod For $2.38M Settlement," Law 360. http://www.law360.com/articles/563583/lexisnexis-retail-workers-get-nod-for-2-38m-settlement.

Krajewski, M. (2011). Paper machines: About cards & catalogs, 1548-1929.

Krogstad, J. M. (2014, March 24). Census Bureau explores new Middle East/North Africa ethnic category. http://www.pewresearch.org/fact-tank/2014/03/24/census-bureau-explores-new-middle-eastnorth-africa-ethnic-category/

Lampland, M., & Star, S. L. (2009). Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life. Retrieved from

Levy, K & Barocas, S ."DESIGNING AGAINST DISCRIMINATION" https://
      scholarship.law.berkeley.edu/btlj/vol32/iss3/5/ doi.org/10.15779/Z38BV79V7K

Liedtke, Michael (2011, Dec 16)   How A Clerical Error Put A Woman On The Streets. Business Insider / AP   . .
      https:// www.businessinsider.com/mistaken-identity-put-this-woman-on-the-streets-2011-12

Majchrzak, A., & Markus, M. L. (2013). Methods for policy research: Taking socially responsible action.

Margetts, H., 6, P., & Hood, C. (2010). Paradoxes of modernization: Unintended consequences of public policy
      reform.

Markham, A. N. (2018). Afterword: Ethics as Impact—Moving From Error-Avoidance and Concept-Driven
      Models to a Future-Oriented Approach. Social Media + Society, 4(3), 2056305118784504. doi:
      10.1177/2056305118784504

Markus, M. L. (2016). Obstacles on the Road to Corporate Data Responsibility. In C. R. Sugimoto, H. R. Ekbia, &
      M. Mattioli (Eds.), Big data is not a monolith (p. 143). MIT Press.

Markus, M. L., & Bui, Q. "Neo." (2012). Going Concerns: The Governance of Interorganizational Coordination
      Hubs. Journal of Management Information Systems, 28(4), 163–198.

Markus, M. L., Steinfield, C. W., & Wigand, R. T. (2006). Industry-Wide Information Systems Standardization as
      Collective Action: The Case of the U.S. Residential Mortgage Industry. MIS Quarterly, 30, 439–465.

Martin, A. (2010). As a hiring filter, credit checks draw questions. New York Times, April 10.
      https://www.nytimes.com/2010/04/10/business/10credit.html

Martin, P. W. (1991). How New Information Technologies Will Change the Way Law Professors Do and Distribute
      Scholarship. Law. Libr. J., 83, 633.

Marwick, A. E., & boyd, danah. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and
      the imagined audience. New Media & Society, 13(1), 114–133.

McArthur, T. (1986). Worlds of reference: Lexicography, learning, and language from the clay tablet to the
      computer. Retrieved from http://www.loc.gov/catdir/description/cam031/85007860.html

McClure, C. R., Hernon, P., & Relyea, H. C. (1989). United States government information policies: Views and
      perspectives. Norwood, N.J.: Ablex.

McKenzie, Patrick (2010) "Falsehoods Programmers Believe About Names"

Meyer, Eugene L. (1997, Dec 15) "Md. Woman Caught in Wrong Net; Data Errors Link Her to Probes, Cost 3
      Jobs," Washington Post C1.

Neff, G. (2012). Venture labor: Work and the burden of risk in innovative industries. Cambridge Mass.: MIT Press.

Neff, G., Tanweer, A., Fiore-Gartland, B., & Osburn, L. (2017). Critique and Contribute: A Practice-Based
      Framework for Improving Critical Data Studies and Data Science. Big Data, 5(2), 85–97. doi:
      10.1089/big.2016.0050

Ostrom, E. (1986). An agenda for the study of institutions. Public Choice, 48(1), 3–25.

Pager, D., & Shepherd, H. (2008). The Sociology of Discrimination: Racial Discrimination in Employment,
      Housing, Credit, and Consumer Markets. Annual Review of Sociology, 34(1), 181–209. doi:
      10.1146/annurev.soc.33.040406.131740

Pasquale, F. A., & Citron, D. K. (2014). Promoting Innovation While Preventing Discrimination: Policy Goals for the Scored Society. Washington Law Review, 89, 1413–1424.

Peeters, R., & Schuilenburg, M. (2018). Machine justice: Governing security through the bureaucracy of algorithms. Information Polity: The International Journal of Government & Democracy in the Information Age, 23(3), 267–280. doi: 10.3233/IP-180074

Pewen, W. F. (2012, August 2). Protecting Our Civil Rights in the Era of Digital Health. The Atlantic. Retrieved from http://www.theatlantic.com/health/archive/2012/08/protecting-our-civil-rights-in-the-era-of-digital-health/260343/

Picchi, Aimee (2019) Consumer Reports: How to Stop Harassment for Debts You Don't Owe. New Economy Project Blog. neweconomynyc.org

Porter, Eduardo. (2018). The hard truths of trying to "save" the rural economy. New York Times.

Powles, J. (2017, December 21). New York City's Bold, Flawed Attempt to Make Algorithms Accountable. The New Yorker.

Raymond, A. H., Young, E. A. S., & Shackelford, S. J. (2018). Building a Better HAL 9000: Algorithms, the Market, and the Need to Prevent the Engraining of Bias. 42.

Robinson, D. G., Yu, H., Zeller, W., & Felten, E. W. (2008). Government Data and the Invisible Hand. Yale Journal of Law and Technology, 11(Spring). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1138083#PaperDownload

Rocher, L., Hendrickx, J. M., & de Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. Nature Communications, 10(1), 3069. doi: 10.1038/s41467-019-10933-3

Rossi, F. (2019). Building Trust in Artificial Intelligence. Journal of International Affairs, 72(1), 127–134.

Ruppert, E. (2009). Becoming Peoples. Journal of Cultural Economy, 2(1–2), 11–31. doi: 10.1080/17530350903063909

Ruppert, E. (2011). Population Objects: Interpassive Subjects. Sociology, 45(2), 218–233. doi: 10.1177/0038038510394027

Schillo, R. S., & Robinson, R. M. (2017). Inclusive Innovation in Developed Countries: The Who, What, Why, and How. Technology Innovation Management Review, 7(7), 34-46.

Schlesinger, A., O'Hara, K. P., & Taylor, A. S. (2018). Let's Talk About Race: Identity, Chatbots, and AI. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, 1–14. https://doi.org/10.1145/3173574.3173889

Singer, N. (2018, September 22). Just Don't Call It Privacy. The New York Times. Retrieved from https://www.nytimes.com/2018/09/22/sunday-review/privacy-hearing-amazon-google.html

Smith, H. J., & Hasnas, J. (1999). Ethics and Information Systems: The Corporate Domain. MIS Quarterly, 23(1), 109–127. doi: 10.2307/249412

Solove, D. (2004). The digital person: Technology and privacy in the Information Age. New York: New York University Press.

Solove, D. J. (2001). Access and Aggregation: Public Records, Privacy and the Constitution. Minnesota Law

Review, 86(6), 1137.

Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. The American Economic Review, 92(3), 434–459.

Star, S. L., & Bowker, G. C. (2007). Enacting silence: Residual categories as a challenge for ethics, information systems, and communication. Ethics and Information Technology, 9(4), 273–280. doi: 10.1007/s10676-007-9141-7

Starr, P. (1992). Social Categories and Claims in the Liberal State. Social Research, 59(2), 263–295.

Stendahl, M., Peter, Walter, B., Carter, & OECD,. (2005). Oslo Manual—Guidelines for Collecting and Interpreting Innovation Data (3rd edition). Organisation for Economic Cooporation and Development, OECD.

Stiglitz, J. E. (1999). On liberty, the right to know, and public discourse: The role of transparency in public life. Oxford Amnesty Lecture. Retrieved from http://www.internationalbudget.org/wp-content/uploads/On-Liberty-the-Right-to-Know-and-Public-Discourse-The-Role-of-Transparency-in-Public-Life.pdf

Taylor, A. G., & Joudrey, D. N. (2009). The organization of information (3rd ed.).

Tempini, N. (2017). Till data do us part: Understanding data-based value creation in data-intensive infrastructures. Information and Organization, 27(4), 191–210. doi: 10.1016/j.infoandorg.2017.08.001

Ticona, Julia, Alexandra Mateescu, and Alex Rosenblat. 2018. Beyond Disruption:How Tech Shapes Labor Across Domestic Work & Ridehailing. New York: Data & Society. https://datasociety.net /output/beyond-disruption/

Tress, W. (2009). Lost Laws: What We Can't Find in the United States Code. Golden Gate University Law Review, 40(2), 129–164.

Ungbha Korn Jenny (2018). Equitable Cities Instead of Smart Cities: Race and Racism Within The Race For Smart Cities. Journal of Civic Media, 1(1), 34-45.

Valentino-DeVries, J., Singer, N., Keller, M. H., & Krolik, A. (2018, December 10). Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret. The New York Times. Retrieved from https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

Walsh, T. (1992). A theory of abstraction. Artificial Intelligence, 57(2–3), 323–389. doi: 10.1016/0004-3702(92)90021-O

Washington, A. (2019). Who Do You Think We Are? The Data Publics in Digital Government Policy. Proceedings of the 52nd Hawaii International Conference on System Sciences. Presented at the Hawaii International Conference on System Sciences, Maui, Hawaii.

Washington, A. L., & Griffith, J. C. (2007). Legislative Information Websites: Designing Beyond Transparency. In A. R. Lodder & L. Mommers (Eds.), Legal Knowledge and Information Systems JURIX 2007 The Twentieth Annual Conference (p. 192).

Washington, A. L., & Morar, D. (2017). Open Government Data and File Formats: Constraints on Collaboration. Proceedings of the 18th Annual International Conference on Digital Government Research, 155–159. doi: 10.1145/3085228.3085232

Washington, A. L., Willis, D., & Tauberer, J. (2012). Do-it-yourself Transparency: Emerging Methods of Congressional Information Dissemination. Proceedings of the 13th Annual International Conference on Digital Government Research, 260–261. doi: 10.1145/2307729.2307774

Weisstein, E. W. (2019, September 6). Birthday Problem. Retrieved September 1, 2019, from MathWorld Wolfram
    website: http://mathworld.wolfram.com/BirthdayProblem.html

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The Role and Limits of Principles in AI Ethics:
    Towards a Focus on Tensions. Proceedings of the 2nd AAAI/ACM Conference on AI, Ethics, and Society, 7.

Winter, S. J., Chudoba, K. M., & Gutek, B. A. (1997). Misplaced resources? Factors associated with computer
    literacy among end-users. Information & Management, 32(1), 29–42. doi: 10.1016/S0378-7206(96)01086-5

Yampolskiy, R. V. (2016). Taxonomy of Pathways to Dangerous AI. 2016 AI, Ethics, and Society Workshop at the
    Thirtieth AAAI Conference on Artificial Intelligence, Pages 143-148. Phoenix, Arizona, USA: AAAI Press,
    Palo Alto, CA.

Acknowledgements

anne.washington@nyu.edu
FSCTestimony Washington SUBMIT2.docx

.