# Safe AI: Practical Policy Approaches

Mariah A. Knowles
Information School
Educational Policy Studies
UW-Madison
baknowles@wisc.edu

Alan Rubel
Information School
Center for Law, Justice, and Society
UW-Madison
arubel@wisc.edu

January 2020

Artificial Intelligence (AI) "matters morally because it motivates large swaths of actors to race to shape its regulation" (Knowles and Rubel 2019). For example, one day an autonomous vehicle will be involved in an accident. Just recently an Uber vehicle struck and killed a jaywalker because the AI did not recognize the jaywalker as a human (Gonzales 2019). So when autonomous vehicles are more widespread, we will need to be able to limit the resulting harms and to redress the harms that still occur. And "AI policy" will play an important role in this mission.

Just banning AI is not the answer. Even though autonomous vehicles will kill some people in accidents, and even if it does so through criminal hacking, their widespread deployment still has promise to lower the total risk of the road. Of course, safer does not mean not perfect; it may be worth treating AI's remaining risks as "unavoidably unsafe" as we already treat the risks inherent in blood banks (*id.*).

Some AI applications, such as consumer robotics, can be "safe" yet still pose new opportunities for old consumer protection issues (Hartzog 2015). Because humans anthropomorphize and establish bonds with our robots, we may be willing to share information with them that we would not share with a human; because these machines will collect data with perfect recall and may always be connected to the manufacturer through the internet, they provide

compelling opportunities for marketers to influence our buying decisions; and because of our attachments to the device, we may be compelled to purchase, for example, a software update under the guise of keeping the robot "healthy" (*id.*).

So, in this review, I consider how the US government could in practice directly regulate AI in order to limit and protect against harms to persons, and I focus on two regulatory approaches often paired together: standards and audits; and torts.

# 1 What is AI Policy?

"Policy" is broader than protecting consumers against harms, and "AI policy" is broader than direct regulation.

There are policies predating AI that scholars have restated in light of AI's recent advancements. For example, Lim (2018) and Ebrahim (2019) compare legal doctrines that would answer questions around AI-generated works, copyright, and patent law. Hallevy (2010) does the same for criminal responsibility when machines accidentally "kill" a human. And Burden (2018) considers implications for contract liability between a business and their AI-using data processors.

Related, there are policies predating AI yet with significant impact on creating, trading, and using AI. For example, Scherer (2017) cautions those developing AI that could help narrow job applicant pools that anti-discrimination laws prohibit both "disparate treatment" and "disparate impact" of people along protected class lines (*eg.*, race, sex). Because "the law prefers to keep employment-related assessments subjective when they involve [protected classes]," AI *as it is today* faces serious challenges in this application area (*id.*).

There are policies and areas of law that are, broadly, "enforced" with AI's aid, such as detecting copyright infringement and reducing patent redundancy. The "AI and Law" literature has researched and developed AI that would aid legal professionals in ensuring consistency with precedent even in the face of complex bases of case law (Horty and Bench-

Capon 2012) and organizing and searching case law relevant to a case at hand (Bench-Capon et al 2012).

There are also policies *incidental* to AI. The proposed Armed Forces Digital Advantage Act (2019) and the proposed Computer Science for All Act (2019) do not seek to increase US education and training in AI specifically, but they do include AI in their definition of computing. The EU's Commission Implementing Decision 2018/321 does not regulate autonomous vehicles directly, but it does seek to regulate a direct resource necessary for autonomous vehicles (high precision GPS data). And the proposed UIGHUR Act (2019), among other things, requests a public report on the nature of certain events in China *because* of China's purported use of AI in large scale and arbitrary policing of groups along protected class lines.

But for this paper, I focus on policy that would (i) *directly* regulate AI's creation, trade, use, and professional training in the US; (ii) have the aim to reduce harms to persons; and (iii) whose regulatory approaches include standards and audits, or torts.

# 2 Definitions

## 2.1 Standards and Audits

Generally by "standards" I mean federal statutes and rules, and by "audits" I mean mechanisms for assessing compliance with those standards. This approach is what Bardach (2015) calls "social" and "protective" regulation. This approach is an attempt to redress the imperfections of relying on the courts due to "poor market information" or other adjudicative "frictions" (*id.*). For example, consider the FDA's role in drug safety. "Scientific uncertainties, technical difficulties of measurement, and political pressures" can complicate relying on this approach (*id.*).

For example, the proposed Artificial Intelligence Initiative Act (2019) and its companion in the House, the GrAITR Act (2019), both, among other things, order the Director of NIST

to meet with AI developers and stakeholders to coordinate the development of standards and measures for assessing AI in relation to security, accountability, and so on. These proposals are concerned with *incentivizing* a *coordinated and standardized* effort by formulating goals and investing in institutions that would advance those goals. Yet, neither puts into place a mechanism for auditing compliance in *existing* AI systems. But the proposed Algorithmic Accountability Act (2019) and its companion of the same name in the Senate would. These proposals would, among other things, require the FTC to define standards for "high-risk information systems" in relation to privacy, bias, and so on; would require data processors to regularly perform self-assessments or third-party audits if able; and would enable the FTC to bring suit for unfair or deceptive practices against data processors who fail to adhere to these standards.

## 2.2 Torts

A tort is a "harmful act (other than a breach of contract) for which courts will provide a remedy... to a private party" (Hall et al 2005). Tort law can be used as a mechanism for allocating risks and liabilities among private parties in an adjudicatively efficient way (Bardach 2015). Tort laws can shift distributions of wealth and risk among parties, such as among consumers and the businesses they interact with (*id.*). To award the plaintiff damages in a tort case, a court must determine that the defendant "breaches a legal duty to [the] plaintiff and this breach proximately causes an injury recognized under law"; but, the definitions and tests of these duties and damages can be diverse across states (Hall et al 2005).

For example, the proposed Algorithmic Accountability Act (2019) would also empower States to bring civil suits on behalf of their residents against data processors who fail to adhere to those FTC standards if that failure has "threatened or adversely affected" any "interest of the residents of [that] State." And the proposed DEEP FAKES Accountability Act (2019), among other things, would define as a new offence the act of distributing harmful

"deep fake" media without appropriate consent of the persons depicted and appropriate disclosure that the media had been algorithmically manipulated. These proposals would also permit civil suits against those who knowingly and with intent commit that offense, with varying levels of damages based on the actual harm faced, the extremeness of the falsely depicted conduct, and any explicit sexual content depicted.

### 2.3 Artificial Intelligence

I use the definition of AI given in the proposed Artificial Intelligence Initiative Act (2019).

## 3 Literature Review

Raw search results are summarized in Table 1. I searched six databases at the combined recommendations of Cat, Alan, the UW Top Law Library list, and the UW Library recommended databases for Political Science, Law, and Philosophy. To my knowledge and based on experience from earlier drafts, these databases covered the literature indexed by other databases recommended by these sources. One common issue was the conflation of "regulation of AI" and "regulation enforced by AI" or so on. After removing works I could immediately deem not relevant by reading titles, abstracts, and conclusions; and after removing cross-database duplicates, I ended with 85 works for potential review. This number fell significantly during closer reads.

| Database | AI | Policy |
|---|---|---|
| **PAIS** | Artifical Intelligence (398) | Regulation + Litigation (20483) |
| Final Results: 13 | * (11866) | ** (162323) |
| **HeinOnline** | Artificial Intelligence (496) | Torts + Regulation (26520) |
| Dates: 2010– | * (59424) | ** (247336) |
| Section Type: Articles | | |
| Final Results: 29 | | |
| **Nexus Uni**[1] | * (1076) | ** (10000+) |

[1]Nexus Uni's final results (just over a thousand) contained duplicate entries. This exacerbated the conflation problem. To deal with Nexus Uni's large raw results, I selected articles for review by skimming for relevant abstracts and titles from the first 100 articles sorted by recency, then again but sorted by relevancy. After removing duplicates, this gave 39 articles for potential review from this database.

| | | |
|---|---|---|
| *(Continued)*<br>Other: "ARTICLE: "<br>Collection: Law Reviews and Journals<br>Category: Law Reviews and Journals<br>Date: 2010–<br>Jurisdiction: Excluding International<br>Practice Area and Topics: Computer and Internet Law: Civil Actions[2]<br>Other: Including Legal Equivalents<br>Final Results: 1067 | | |
| **Index to Legal Periodicals**<br>Limit To: Scholarly<br>Final Results: 45 | Artificial Intelligence + Artificial Intelligence Laws (594) | Torts + Government Regulation + Artificial Intelligence Laws (34421) |
| **Phil Index**<br>Language: English<br>Source Type: Journals<br>Final Results: 2 | Artificial Intelligence (674)<br>* (2825) | Regulation + Litigation + Tort (1217)<br>** (13937) |
| **Phil Papers**[3]<br><br>Restrictions: Published Only<br>Final Results: 39 | "artificial intelligence" + "autonomous vehicles" + "deep fakes" + "robot" (974) | "regulation" + "torts" + "private law" + "public policy" (998) |
| | * "artificial intelligence" + "AI" + "machine learning" + "machine-learning" + "big data" + "autonomous vehicles" + "deepfakes" + "deep fakes" "deep-fakes" + "robot" + "algorithm" + "predictive analytics" + "predictive policing" + "data mining" + "data-mining" | ** "FTC" + "regulation" + "assessment" + "evaluation" + "measurement" + "audit" + "standards" + "NIST" + "civil liability" + "tort" + "damages" + "civil suit" + "litigation" + "injury" + "Federal Trade Commission" + "FCC" + "Federal Communications Commission" |

Table 1: Raw literature search results. Controlled vocabulary terms are shown as Proper Nouns, keyword terms are shown in "quotes," and intermediate number of results for each term are shown in (parentheses). Disjunction within a term is shown with a +plus. The Final Results for each database refers to the conjunction of all search terms for that database. If a date range or other filter is noted for a database, then the final result and intermediate result counts for that database all reflect that filter.

# 4 Analysis

Across recommendations in the literature, federal agencies (most notably the FTC) played an important role in protecting consumers from AI-attributable harms. There were counter-recommendations, however, that saw the federal government as limited in its ability to keep

[2]This option could not be selected for the Policy-only search

[3]Phil Papers had very short limits on number of query terms. This database also provided no interface for searching in the intersection of two categories. So, I relied on the category hierarchy to inform my terms.

up with AI's rapid development and instead recommended industry self-regulation. Regardless, state courts were recommended as the body to determine liability in any particular case after-the-fact, since apparent AI-attributable harm may come with complex and unclear liability lines. Finally, many recommendations considered the role of insurance as influencing parties before-the-fact and footing the bill after-the-fact.

## 4.1 Federal Agencies

Tort law alone may be too slow to respond to technological change and bad actors' inventive ways to use AI deceptively (Hartzog 2015). The FTC could take the lead in protecting consumers from unfair and deceptive practices attributable to AI, since it already has a long history of dealing with these harms in other domains (Hartzog 2015; Tschider 2018; Hirsh 2014). However, the FTC may need to work with other agencies, such as the FDA involving cases of mechanical implants and the NHTSA involving cases of autonomous vehicles (*id.*).

Existing agencies may face difficulties keeping abreast of technology changes while also keeping up with their current duties and tight budgets (Hartzog 2015). Policymakers could ease this burder by establishing a new agency that could advise other agencies on matters of AI (*id.*).

Whether existing, new, or some combination, agencies could define safety standards for AI in various domains. They could certify AI systems, granting the owner of the system some form of shield against legal liability (Scherer 2016). They could define the operational domains of various AI and encourage governments to pass regulations to prevent "dead zones" of external resources necessary to those systems, such as GPS data for autonomous vehicles (Cowger 2018).

As with many types of products, autonomous vehicle manufacturers, for example, will be required by federal and state agencies to disclose inherent risks to their passengers. Disclosing this risk unambiguously and in a way that could be compared across the market may be difficult given the evolving complexities of AI technologies. One recommendation

is to disclose the insurance premium for the vehicle's fleet as a way to quantify that risk (Geistfeld 2017). And though it varies from state to state, manufacturers will have a duty, underlined by federal standards, to prevent manufacturing defects and design systems that are safe when functioning properly (*id.*).

A concern throughout much of the AI literature is the "Trolley Problem" when an AI system, *without* defect and *with* safety features, will be in a situation where it must decide who to harm and will not have a human to defer to. Consider autonomous vehicles when an accident is imminent. In these cases, the AI should treat inhabitants and bystanders involved equally (Geistfeld 2017). And agencies should limit the types of data an AI may consider in that decision making and should ban alternative algorithms which would, in the event of an imminent accident, default to harm persons based on their apparent race, religion, or so on (Cowger 2018).

The more an autonomous vehicle drives, the more data it collects, and conceivably the safer it will be (Geistfeld 2017). But we can also pool the miles of vehicles in a "fleet" of like model and operating system, since their data can be shared between them (*id.*). Federal agencies could then set minimum requirements for the number of pre-market miles driven by a manufacturer's fleet (*id.*). Federal rules may also need to specify optimal safety standards for autonomous vehicles (*id.*). If federal standards are not optimal, but instead minimal, then state-level safety regulations may pre-empt the federal-level standards, which will cause greater legal uncertainty for manufacturers and insurers (*id.*). This uncertainty could delay deployment for technologies that have promise to improve public safety.

When federal agencies assess the safety of autonomous vehicle fleets, it may be appropriate to compare the accident rates of the fleet versus humans in similar situations (Geistfeld 2017). Assessments should not be limited to *lethal* accident rates, since lethal accidents are rare, and so the number of miles of testing required by manufacturers to show statistically significant improvement would skyrocket (*id.*). Again, this may cause unnecessary delays.

The power of agencies to regulate AI has its complications. First, the question is unset-

tled: should we be comparing AI systems' promises of safety to our current non-AI systems as they are now, or to the possible human systems as they in the future when "assisted" by AI safety features (Calo 2017)? Second, although federal standards and audits have better before-the-fact regulatory power than the courts with emerging technology, "[i]t may well be that additional research, development, and even public operation are the only ways to determine which types of AI are harmful and which types are not" (Scherer 2016). Third, agencies operate under restrictions that may slow their efforts to add or change rules: by the time a new rule could take effect it may be too late to prevent the targeted harm, or the targeted class of AI may already be obsolete (*id.*). Fourth, reliance on agencies can upset voters who see agency actors as unelected and therefore undemocratic (*id.*). And finally, although agencies do provide governments a level of expertise, what "expertise" means is unclear and what expertise they provide can wane as technology changes; when that happens, it can again be the case that only researchers directly involved with the technology's development will retain the experience necessary to make accurate safety assessments (*id.*).

## 4.2 Industry Self-Regulation

Those complications lead some in the literature to instead recommend self-regulation within the AI industry (such as Google, research institutions, and so on), but "nudged" by the government. For example, a government could certify AI systems as a shield against liability, but it could just as well be an industry body doing the certifying, with only limited oversight from the government into that process (Yanisky-Rawid and Hallisey 2019). It is the major industry actors that "hold the lion's share of regulatory resources" needed to understand and police AI before-the-fact; and when we are facing existential threats from AI, after-the-fact reliance on the courts cannot by definition work (Guihot et al 2017).

This does not mean the government should stay out of the regulatory picture. Governments could "nudge" or "sway" industry self-regulation by signaling their intentions and investing in initiatives consistent with those missions (Guihot et al 2017). Legislatures, for

example, could propose bills that signal (i) the government's intentions regarding issues of technology, society, and the economy and (ii) the government's willingness to regulate activities related to these issues (*id.*). And, "[b]y publishing industry standards and engaging in public advocacy, agencies can set expectations without specifically approving or banning particular products," which, as noted above, could face counterproductive delays (Scherer 2016). And there are currently efforts using this strategy, such as with the Partnership on AI, AAAI, and IEEE (Guihot et al 2017). However, governments can "retreat" from these partnerships and nudges, which can effectively send an inverse signal to the industry (*id.*).

The industry could self-regulate through ethics review boards and/or codes of ethics. But there is high variability in the application, technology, and associated risks of the AI that could come under review, and these ethics may be too high level or misaligned with that high variability (Guihot et al 2017). Further, we might not be able to rely on the industry to self-regulate with codes of ethics because the FTC has brought suits more than once against industries whose codes were seen as limiting trade (Calo 2017). Also, any type of industry (self-)regulation is complicated by trade secret laws, which could inhibit sufficient assessment of AI technologies; therefore, to counter these complications while still preventing assessments from "outing" trade secrets, whistle-blower protections may need to be strengthened and relied upon in any AI regulatory strategy (Katyal 2019; Desai and Kroll 2017).

Finally, an industry self-regulation approach does not have to happen in isolation. By offloading much of the complexity to the private sector, the federal agencies could focus on issues of the highest risk to the public while still preparing the agency's future actions to be responsive to technological change (Guihot et al 2017).

## 4.3 Courts

While AI standards (whether federal or industry) would aim to reduce public risks, some AI-attributable harm is inevitable. The role of the courts after-the-fact will be to determine

who should be liable and to decide which test(s) should be used. Variously, the literature suggested strict liability doctrines, where the plaintiff only needs to show that the harm was proximately caused by the product; and negligence liability doctrines, where the plaintiff would also need to show that the defendant failed in its duty to ensure safety.

The question of who to hold liable is at least simplified in the case of autonomous vehicles because many manufacturers of autonomous vehicles have elected to assume liability (Geistfeld 2017). They benefit from this because it allows them to avoid "a hodge-podge of regulations all aimed at reducing or allocating damages," and they could even attract customers to use that manufacturer's vehicles if that assumption of liability could be spun as a selling point (Cowger 2018).

The next question is how to test that liability. When the harm is attributable to a failed piece of hardware or an intentional maliciousness directly attributable to one human, then the courts can rely on a wealth of product liability precedent (Geistfeld 2017; Cowger 2018). The "new" considerations in light of AI, setting aside the question of whether anything is ever "new" in law, include two plausible situations identified by the literature. First, when an autonomous vehicle's AI operates properly and yet still must choose who to harm, we may be able to rely upon strict liability doctrine coupled with safe harbors for those compliant with federal or industry standards (Geistfeld 2017; Cowger 2018; Yanisky-Rawid and Hallisey 2019). And second, when an autonomous vehicle's AI is hacked and used by terrorists to cause harm, the appropriate doctrine was left unclear besides a tentative suggestion to rely upon negligent liability and make a judgement of reasonable care exercised by the manufacturer during its pre-market testing (Geistfeld 2017).

However, courts cannot be relied upon to resolve all after-the-fact harms of AI. Courts might only see cases on the most "visible" of AI technologies that juries can understand, like the autonomous vehicle examples throughout (Scherer 2016). And, technological arguments can be speciously made in courts, since juries may be more susceptible to misleading information (*id.*). This exacerbates the need for official bodies (whether federal or industry) that

can be better informed and can better focus on broader, and less "visible," public risks.

### 4.4 Insurance

Finally, after the court's determination, insurance will likely be relied upon to distribute the costs over the entire industry (Cowger 2018). And before-the-fact, the insurance sector can even influence safety standards for the better. In fact, "scholars have often celebrated insurer's conduct-forcing capabilities in the commercial space, ascribing improvements in product and worker safety" (Helveston 2016). One recommendation considers insuring AI systems similar to how we currently insure vaccinations (Abbot and Sarch 2019).

However, manufacturers will not be the only insured parties, and consumer insurance companies may also be using AI in making decisions such as whether to insure or how high to set premiums (Helveston 2016). This, if left unregulated, can have an unfair influence over private persons' lives, since: getting insurance can carry the power of government mandate; the market of available insurers can be sparse; and unchecked AI systems might determine one's insurance premium based on, in part, for example, that person's apparent religion on social media (*id.*). Therefore, one recommendation is to restrict what characteristics can be used at all in making these types of decisions, similar to the community rating mandates of the Affordable Care Act (*id.*).

## 5 Conclusion

The same strategies for regulating AI are found in the letter of proposed law and in the reviewed literature. This literature points to the importance of each regulatory body's role in the broader picture, and this literature spells out various limitations that these proposed bills may face.

Most importantly, it is clear—even within this purposefully limited search—that regulating AI cannot be reduced to two complementary approaches alone. One can find evidence

of any of Bardach's (2015) example approaches to regulation in use, in proposals, and in the literature, all with respect to AI. And although AI "motivates large swaths of actors to race to shape its regulation" (Knowles and Rubel 2019), it is clear that each of those actors may be composed of complex networks of interacting regulatory bodies, each with its own roles and precedent to draw upon.

# Bibliography

Abbott, Ryan and Alex Sarch. "Punishing Artificial Intelligence: Legal Fiction or Science Fiction." *UC Davis Law Review* 53. 2019.

Bardoch, Eugene. *A Practical Guide for Policy Analysis.* 2015.

Bench-Capon, Trevor et al. "A History of AI and Law in 50 Papers: 25 Years of the International Conference on AI and Law." *Artificial Intelligence and Law* 20(3). 2012.

Burden, Kit. "Impact of Disruptive Technologies on Sourcing and Outsourcing Transactions." *Computer Law and Security Review* 34(4). 2018.

Calo, Ryan. "Artificial Intelligence Policy: A Primer and Roadmap." *UC Davis Law Review* 51(2). 2017.

"Commission Implementing Decision 2018/321." *Official Journal* L62(34). 2018.

Cowger, Alfred. "Liability Considerations when Autonomous Vehicles Choose the Accident Victim." *Journal of High Technology Law* 19. 2018.

Desai, Deven and Joshua Kroll. "Trust but Verify: A Guide to Algorithms and the Law." *Harvard Journal of Law and Technology* 31. 2017.

Ebrahim, Tabrez. "Data-Centric Technologies: Patent and Copyright Doctrinal Disruptions." *Nova Law Review* 43. 2019.

Geistfeld, Mark. "A Roadmap for Autonomous Vehicles: State Tort Liability, Automobile Insurance, and Federal Safety Regulation." *California Law Review* 105(6). 2017.

Gonzales, Richard. "Feds Say Self-Driving Uber SUV did not Recognize Jaywalking Pedes-

trian in Fatal Crash." *NPR*. 2019.

Guihot, Michael et al. "Nudging Robots: Innovative Solutions to Regulate Artificial Intelligence." *Vanderbilt Journal of Entertainment and Technology Law* 20. 2017.

Hall, Ketmit et al. *The Oxford Companion to the Supreme Court of the United States.* 2005.

Hallevy, Gabriel. "Virtual Criminal Responsibility." *Original Law Review* 6(1). 2010.

Hartzog, Woodrow. "Focus on Cyberlaw: Unfair and Deceptive Robots." *Maryland Law Review* 74. 2015.

Helveston, Max. "Consumer Protection in the Age of Big Data." *Washington University Law Review* 93. 2016.

Hirsch, Dennis. "That's Unfair! Or is it? Big Data, Discrimination and the FTC's Unfairness Authority." *Kentucky Law Journal* 104. 2014.

Horty, John and Trevor Bench-Capon. "A Factor-Based Definition of Precedential Constraint." *Artificial Intelligence and Law* 20(2). 2012.

Katyal, Sonia. "Private Accountability in the Age of Artificial Intelligence." *UCLA Law Review* 66. 2019.

Knowles, Bryan and Alan Rubel. "Eight Views Reviewed: A Review of the Literature of Why AI Matters Morally." *Preprint.* 2019.

Lim, Daryl. "AI and IP: Innovation and Creativity in an Age of Accelerated Change." *Akron Law Review* 52(3). 2018.

Scherer, Matthew. "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies." *Harvard Journal of Law and Technology* 29. 2016.

Scherer, Matthew. "AI in HR: Civil Rights Implications of Employers' Use of Artificial Intelligence and Big Data." *The SciTech Lawyer* 13(2). 2017.

Tschider, Charlotte. "Regulating the Internet of Things: Discrimination, Privacy, and Cybersecurity in the Artificial Intelligence Age." *Denver Law Review* 96. 2018.

United States. Cong. House. Algorithmic Accountability Act of 2019. 116th Cong. HR 2231.

United States. Cong. House. Computer Science for All Act of 2019. 116th Cong. HR 1485.

United States. Cong. House. DEEP FAKES Accountability Act. 116th Cong. HR 3230.

United States. Cong. House. GrAITR Act. 116th Cong. HR 2202.

United States. Cong. House. UIGHUR Act of 2019. 116th Cong. HR 1025.

United States. Cong. Senate. Algorithmic Accountability Act of 2019. 116th Cong. S 1108.

United States. Cong. Senate. Armed Forces Digital Advantage Act. 116th Cong. S 1471.

United States. Cong. Senate. Artificial Intelligence Initiative Act. 116th Cong. S 1558.

Yanisky-Ravid, Schomit and Sean Hallisey. "Equality and Privacy by Design: A New Model of Artificial Intelligence Data Transparency via Auditing, Certification, and Safe Harbor Regimes." *Fordham Urban Law Journal* 46. 2019.

.

# Process Comments

Jan 3rd—Met with Cat. Discussed my experiences doing literature reviews. Discussed ways to improve upon original MD1 search methods. In particular, discussed controlled vocabulary and ways to display raw results.

Jan 3rd–8th—First draft of revised search. Searched Digital National Security Archive; EUR-Lex; HeinOnline; National Journal; Wisconsin Digital Archives; WisconsinEye; Index to Legal Periodicals and Books; Nexus Uni; ProQuest Congressional; Wisconsin Legislative Drafting Records; Wolters Kluwer Cheetah; and WorldCat.

Jan 8th—Met with Cat over Skype. Discussed a draft of an initial set of search results. Discussed ways to better improve how I find and select databases. In particular, discussed the UW Library Database subject listings.

Jan 8th–21st—Second draft of revised search and substantive writing. Searched PAIS; HeinOnline; Nexus Uni; Index to Legal Periodicals; Phil Index; and Phil Papers.

Jan 21st—Completed full draft. Sent to Alan.

Jan 22nd—Met with Alan. Discussed full draft. Discussed improvements that still needed

to happen. In particular, discussed making explicit the focus on protecting against harm (vs. focus on AI policy more broadly) and continuing to restructure around themes not pieces.

Jan 26th—Completed revised draft.

Jan 30th—Met with Writing Center. Read aloud section 4. Discussed style and clarity. In particular, discussed specific sentences that should be broken up, specific awkward phrases, and specific citations that needed to be cleared up.

Jan 31st—Completed final draft. Sent to committee.