Regulation of Artificial Intelligence in US Internet Law

A Research Paper
in STS 4600
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By

Matthew Schaeffer

March 14, 2018

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis Related Assignments.

Signed: _____   Date: _____

Approved: _____   Date: _____
Rosalyn W. Berne, Associate Professor, Department of Engineering and Society

Abstract

Artificial Intelligence (AI) is a large technological industry that lacks a proper legal definition. The complexities and issues that arise as a result of this vagueness surrounding AI have only gotten worse over time. As a result, developing and passing proper legislation to regulate companies in the field is seemingly impossible. The objective of this paper is to explore and propose possible regulatory measures that can be enacted upon the online AI industry so that these technologies continue to benefit society without infringing human rights and freedoms. The question is how the industry of AI, particularly those designed for use on the internet, can be regulated to ensure human rights are upheld without stunting the development of the industry. As AI is still a budding industry with the US leading the way, it would be detrimental to enact specific regulation at the current time until further research is done. Creating general laws to govern company and research ethics and providing incentives to companies to promote ethical practices in their AI creation are much more practical solutions than banning certain forms of research and development.

15

Introduction

Many scientists, developers, and tech companies believe that there is no way to regulate AI without a serious analysis that weighs the potential threat AI systems pose against their benefits (Markoff, 2016). AI has the ability to perform great good through stimulating human interaction, recognizing patterns at a far greater rate than humans, and developing new and unique solutions to existing problems in virtually every industry. The big issue with AI is that there are virtually no restrictions on how it can be applied. AI has already been developed that can accurately discriminate certain ethnic groups and even sexualities better than humans can, meaning the potential for misuse is enormous. Recently, a software was created that is able to distinguish gay from straight men at an accuracy of 91%, and women at 83% when given 5 facial images, compared to 61% and 54% respectively from human judges (Wang & Kosinski, 2017). The research topic will focus on how AI can be regulated so that these dangers cannot be exploited for personal or authoritarian means. The safety and privacy of individuals must be paramount; however, the AI industry must remain free enough to explore these often uncomfortable truths without paving the way for abuse and tyranny.

In order to address this potential for great evil, researchers, engineers, and scientists believe that Artificial Intelligence must be regulated and acted upon soon (Müller and Bostrom, 2014). Studying the regulation of AI using Actor-Network Theory (ANT), the biggest actants in the field are the major technological corporations such as Google, Microsoft, Facebook, Apple, and Amazon. The major actors are the leaders of these companies as well as select professors and lawyers from leading universities like Carnegie Melon University (CMU) and Stanford University. In the near future, it's expected that Congress and lawmakers will play vital roles as actors by creating the laws or even a new agency to deal with the concerns over abuse of AI

(Sherer, 2015). By understanding these actants' and actors' roles in regulating AI and their biases, it becomes apparent as to how these laws and regulations may be devised and applied across all AI systems and companies.

Human Rights and Technology

Arguably, the two most important documents for human rights in the modern era are the United States' Bill of Rights and the United Nations' Universal Declaration of Human Rights. These rights are inalienable and are protected from being taken away from the people by a national or international government or any other entity (United Nations General Assembly, 1948). Specifically, the Preamble of the Declaration of Human Rights and the First Amendment of the Bill of Rights state that the right to freedom of speech, press, and expression shall not be compromised by any law. With regards to technology, the United States Supreme Court has ruled that social media are protected platforms for lawful speech, and any laws infringing upon that right are unconstitutional, even in the case of felons of sex crimes (Packingham v. North Carolina, 2017). This 8-0 ruling (Justice Gorsuch abstained) suggests that access to a social media platform is equivalent to having access to a public forum, and thus the right of a person to access such sites cannot be encroached upon in the United States.

Traditionally, the Constitution has only restricted the government in this regard and allowed private companies freedom to do as they please. This has developed into a bit of an issue when it is clear that several of the large technology companies possess greater power and aptitude to censor than the federal government itself. This is especially true when it's worth considering that these companies' market value are all directly tied to the information business. Usually, this would be met with the free market, and alternative technologies where censorship

does not take place would simply be introduced and preferred. In this particular case however, these companies control the ability to identify and verify their own competition. Alternatives cannot be explored if they are not permitted to be viewed by these larger companies in the first place. Regulatory measures must be put in place to prevent the censorship and restriction of human rights by these large companies, particularly with regard to artificial intelligence.

The Stanford Project and AI Development

The most prevalent study in regulation of Artificial Intelligence was conducted by a panel of experts brought together by Stanford University (Stone et al., 2016). They argue that regulation of the industry is not as crucial as the public believes it to be, given the rate of progress. By 2030, they believe that the impact of AI will be almost entirely positive, and have a great and far sustaining impact on our very culture and way of living. The Stanford Project warns that, "if society approaches [AI] primarily with fear and suspicion, missteps that slow AI's development or drive it underground will result, impeding important work on ensuring the safety and reliability of AI technologies" (p. 5). Of course, their input is likely biased given that the panelists are all leaders in the private sector of the AI industry, and as such would love to continue their research unimpeded by heavy regulation. The freedom to explore any variety of applications for AI means that anyone can develop and create a startup so long as there is demand for the research or software. This is already noticeable in places like Silicon Valley in California, which has propelled the state to becoming the 6th largest economy in terms of overall GDP in the world over countries such as France, India, and South Korea (U.S. Bureau of Economic Analysis, 2016).

A paper from the Harvard Journal of Law and Technology proposes a possible method to

regulate future AI (Sherer, 2015). Sherer proposes what he calls the "Artificial Intelligence Development Act." Its mission is "to ensure that AI is safe, secure, susceptible to human control, and aligned with human interests, both by deterring the creation of AI that lack those features and by encouraging the development of beneficial AI that include those features" (pp. 394-396). The act would also establish a government agency similar to the FDA that would approve AI on the basis of their established safety regulations. Instead of forcing all AI to comply, which would be bureaucratic and slow and stifle the innovation of AI, Sherer suggests that, much like the FDA, AI can be sold without the certificate of safety, but the manufacturer will not enjoy the "limited tort liability" that the certificate protects manufacturers from (pp. 395). Certified AI would enjoy the luxury in court of the plaintiff needing to establish actual negligence in the design, manufacturing, or operation of an AI system in order to prevail on a tort claim. On the other hand, non-certified AI would leave the manufacturer totally liable to any injuries or claims. This method would heavily encourage non-certifiable AI to maintain safety precautions, while giving the duality of suggesting AI to be certified by the government agency.

A study about the positive and negative effects of Artificial Intelligence in our lives discussed how it was likely that AI would bring us closer together and stimulate us, rather than making us lazy and unproductive (Gillies, 1996). Gillies believes that regulation would not be necessary until it hurts overall production and intellectual stimulation. Any form of regulation in the market could otherwise stifle competition and progress within the field, hurting any potential AI has for success.

Potential Threats to Human Rights

Another study discusses the overall potential for regulation of this vague and all-

encompassing field (Markoff, 2016). A group of experts from five major tech corporations, Facebook, Amazon, Apple, Google, and IBM, believe that an "ethics panel" within the private sector must be created to ensure the technology is not misused. They believe that it must be done sooner rather than later, as AI is developing quickly. The issue with this is that it is essentially allowing for collusion between major companies to consolidate their power and would likely be subject to anti-trust laws. While we cannot be certain that these companies would abuse the system in that way, the creation of such an exclusionary panel would pose a major threat to consumers and smaller companies.

With the interconnection of people worldwide through the internet, these companies have more influence on thought and greater access to privacy than any other group of corporations in the history of the world. They have developed products that are integral to our everyday lives and are trusted with incredible amounts of personal information. These companies are also the same ones that are leading the development of new AI. Given all of the data these companies have access to, the potential for abuse through blackmail, extortion, invasion of privacy, and even discrimination based on race, sexuality, gender, and thought is extraordinary. This would extend not only to private citizens, but also public and elected officials. In particular, Google is in the process of developing an online AI known as Perspective whose mission is to "make it easier to host better conversations" words (Hosseini, Kannan, Zhang, & Poovendran 2017).It does this by rating comments on a score of "toxicity," an arbitrary figure that can be used to filter comments based on certain phrases or words (Hosseini et al., 2017). It's also consistently possible to deceive the Perspective AI, according to Hosseini et al., by using deliberate misspellings and sarcasm. This suggests that such an AI would have little effect to curb "trolling and harassment," as it states in its objective, but would instead provide the foundation for

censorship on a massive scale (pp. 2). While this particular AI is being used to curb hate speech and harassment online in comment sections such as those of the New York Times and The Guardian, the potential for misuse and censorship is huge. There are no real laws in place that could prevent this level of censorship.

Google is a private company, meaning that the Bill of Rights and other such documents regarding the government do not affect it. However, there is still concern that such large corporations may violate human rights. There is no competition to contest Google; it owns almost all of the search engine market and most of the video market through Youtube (Patterson, 2013). Having control over these particular industries means that even if there were competition, it is possible nobody would know or have access to it because Google essentially controls what is seen and heard on the internet. Patterson (2013) contests that market share, while meaningful, is not the most important criteria to evaluate these online companies on. It is much easier to find alternatives online than it is in the actual market; all it takes is one click away to a competitor, and thus the market barriers for smaller companies are much smaller. The primary concern should therefore be that Google's quality of search results means that consumers remain ignorant to potential alternatives, and often consumers won't attempt to even look for alternatives. This allows Google to intentionally provide weaker quality with their results should they choose to. Patterson (2013) also notes that Google used to claim objectivity in search results in 2007, but has since switched to saying search results are the "opinions" of the company and are biased toward what the company believes to be most useful to consumers (pp. 22). Combined with its AI that can effectively censor criticisms, other companies, and even people, Google, Facebook, Twitter, and Amazon have total control over virtually all information. Such censorship would certainly constitute human rights violations to freedom of speech and expression. It is therefore

concerning that a company that has more influence on human development than almost any individual country in the world would be immune from being punished for it simply on the basis that it is not a nation-state.

Additionally, Google is used as a revenue model for many companies (Patterson, 2013). Almost all of web traffic is directed by Google and as a result, the company can censor some companies and give preference to those that pay Google. This is already seen through its AdWords system and News search which is very selective as to which publications it links to. Even if the information exists on the web, it can oftentimes be impossible to find it in the sea of irrelevancy that these major corporations create.

Discussion

There can be no doubt that Google and other large tech companies essentially control the flow of information around the world. These same companies also are the biggest developers of AI, which is already proving to become a great threat to freedom of speech and privacy online. AI can easily amass large amounts of data and find patterns far quicker than any human could; thus, restricting what information AI have access to is one of the easiest ways to decrease their potential for harm (Markoff, 2016). For example, restricting classified and personal information from AI would greatly protect the American consumer from unwanted rights violations.

None of this matters, however, if the AI is out of control from its creator. Several ways to craft AI to be friendlier include engineering the capability to acquire basic ethics knowledge within the AI, providing stable hierarchical goals, ensuring the early stages of the AI focus on recursive self-improvement with rich human involvement, and promoting the quick development of AI rather than stifling it (Goertzel and Pitt, 2012). These basic tenants would at the very least

keep AI from "going rogue" and force manufacturers to accept responsibility for their creations. If these rules are encouraged by the government for AI companies to follow, the liability for AI acting on its own would lie solely on the developer. On the other hand, should AI be intentionally used maliciously, such uses would fall to the sole liability of the user. For example, suppose an AI were developed to sweep metadata of purchases of a product. If the AI happens to obtain information it should not be privy to, the manufacturer would be liable. However, if somebody were to use this AI to track and target a specific individual, that responsibility would fall on the user. This is very similar to how cars are regulated. An accident caused by a defect falls on the car manufacturer, while accidents or malicious uses of the car fall entirely on the driver. This would protect companies from users abusing their systems while at the same time encouraging companies to develop only useful and safe AI. If the goal is to protect individuals from the harm that can be inflicted by AI, then the government should create laws that reflect the tenants suggested in Goertzel and Pitt's (2012) article. By following these suggestions, the biggest issue becomes not the AI and developers, but rather the users and potential for misuse.

Conclusion

The primary goal of regulating artificial intelligence should not be to confine the industry to specific standards and research, but rather to protect the individual from potential rights violations and threats. The greatest threats will not come from the AI itself. Instead, the greatest potential for harm will come from misuse of AI by users to target specific individuals. General regulations governing what types of information can be accessible to AI are important, as well as laws that promote ethical practices and research. Further regulation could result in decreasing the variety of companies and systems that are developed, stifling innovation and harming the

consumer. The biggest issues lie in who gets to develop these AI, not so much as what types of AI are developed. The fewer the companies in the industry, the smaller the diversity of ideas, and the greater the power concentrated in each company. All of this would lead to more chances for the development and abuse of dangerous AI.

References

Domingos, P. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World (New York, New York: Basic Books, 2015).

Gillies, D. (1996). Artificial intelligence and scientific method. Oxford: Oxford University Press.

Goertzel, Ben and Pitt, Joel (February 2012). Nine Ways to Bias Open-Source AGI Toward Friendliness. Journal of Evolution and Technology **22** (1): 116-131. ISSN 1541-0099.

Hosseini, H., Kannan, S., Zhang, B., & Poovendran, R. (February 2017). Deceiving Google's Perspective API Built for Detecting Toxic Comments. Network Security Lab, Department of Electrical Engineering at the University of Washington.

Markoff, J. (2016, September 01). How Tech Giants Are Devising Real Ethics for Artificial Intelligence. *The New York Times.* Retrieved from https://www.nytimes.com

Müller, V., and Bostrom, N. (2014). Future progress in artificial intelligence: A Survey of Expert Opinion. *Fundamental Issues of Artificial Intelligence*.

Packingham v. North Carolina. 582 U.S. (2017).

Patterson, Mark. (July 2013). Google and Search Engine Market Power. *Harvard Journal of Law and Technology*. Retrieved from http://jolt.law.harvard.edu/assets/misc/Patterson.pdf

Russell, S. and Norvig, P. Artificial Intelligence: A Modern Approach (3rd Edition). Essex, England: Pearson, 2009.

Scherer, Matthew U., Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies (May 30, 2015). *Harvard Journal of Law & Technology*, Vol. 29, No. 2, Spring 2016.

Shaw, M. (2002). Prospects for an Engineering Discipline of Software. *Encyclopedia of Software Engineering*. doi:10.1002/0471028959.sof259

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., … and Teller, A. (2016,

September). Artificial Intelligence and Life in 2030. One Hundred Year Study on

Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University,

Stanford, CA. Doc: http://ai100.stanford.edu/2016-report.

United Nations General Assembly. (December 1948). The Universal Declaration of Human

Rights. https://www.un.org/en/universal-declaration-human-rights/

U.S. Bureau of Economic Analysis. (2016). Gross State Product in California.

http://www.dof.ca.gov/Forecasting/Economics/Indicators/Gross_State_Product/

Wang, Y. & Kosinski, M. (2017). Deep Neural Networks are More Accurate than Humans at

Detecting Sexual Orientation from Facial Images. *Journal of Personality and Social

Psychology*. American Psychological Association.

Yampolskiy, R. V., & Fox, J. (2012). Artificial General Intelligence and the Human Mental

Model. *The Frontiers Collection Singularity Hypotheses,* 129-145. doi:10.1007/978-3-

642-32560-1_7