MEMORANDUM FOR ACTION

TO:	**The Minister of Foreign Affairs**
CC:	**The Digital Inclusion Lab, Office of Human Rights Freedoms and Inclusion**
SUBJECT:	Gendered and Racial Bias in Artificial Intelligence

SUMMARY:

1.   Artificial Intelligence (AI), by design, embeds pre-existing social values into its decisions. With the direction that the technology has taken so far, AI could potentially violate Article 2 of the Universal Declaration of Human Rights by increasing gender and race discrimination. Lack of regulation around innovation risks exacerbating inequalities by replicating unfair biases. When governments become implicated in a potentially harmful technology, it is their responsibility to ensure that these biased and discriminatory practices are addressed, and that corrective measures are taken. Therefore, in setting an example for the world, Canada could potentially help eliminate gender and racially biased AI and ensure its global position as a fighter for social justice. Biases enter AI in different ways. They can be embedded through training data, and also from within the development teams, which are often homogenous. The implications of under-regulated AI could be serious, which is why inclusive hiring practices as well as diverse training data are imperative to ensuring the standardization of fairness in algorithmic technologies. It is important that the technology industry considers taking action, such as implementing inclusion riders, in their hiring policies. AI is a fast-evolving technology, which is why corrective measures should be made with a sense of urgency. The following recommendations are placed in order of feasibility.

RECOMMENDATION(S):

- To encourage that AI is developed in an ethical and fair manner, we ask that **diversity** in training data as well as developer teams should be a priority, and consulting with various disciplines should be emphasized.
- To consider **algorithmic impact assessments** in the design **review processes** for new software as well as its consequent updates.
- We recommend greater regulatory attention to bias in AI to **ensure fairness** and standardization of bias reviews as part of the AI production process.

[Aurelia Talvela, Courtney Blamey, Nina Morena]
[Department of Communication Studies, Concordia University]

☐ I wish to discuss
☐ I concur   ☐ I do not concur

_____

Minister

BACKGROUND:

2.   Artificial Intelligence (AI) mirrors the status quo, and it contains bias. These biases are problematic because they often misrepresent marginalized groups, often through negative gender and race stereotypes. Our research scan has found that AI could potentially violate human rights by increasing discrimination. There have been many accounts of the harms that biases and discrimination in AI can cause. Examples vary from algorithms used in the criminal justice system being biased against African-Americans to algorithms determining job qualifications discriminating against those who have foreign-sounding names. Digital assistant Siri's inability to provide quick and helpful advice to those experiencing sexual abuse, for example, allows this technology to reside in patriarchal values, which might not find this a design priority. Divorce artificial intelligence software is an example of LegalTech, which, while offering a democratizing promise to provide marginalized or otherwise vulnerable groups with access to the law, in fact creates the conditions for unfair divorce agreements to pass due to the lack of a human mediator. In short, biased AI can cause significant disruption in someone's life as it only perpetuates their already marginalized position in society.

3.   Lack of greater regulatory scrutiny in innovation is reckless as it will only serve to exacerbate inequalities by replicating unfair biases. Various biases from the social world are merely being replicated in AI because of value-laden training data, while internal reviews of design processes are also lacking. These values most often reflect social inequalities that have historically marginalized women and people of color. In short, social inequalities are being reiterated and reproduced through training data sets in machine learning. Algorithms have no capacity for honesty. They can only work with the information that is provided and are not capable of understanding historical or social context. In other words, the machine cannot teach; it can only work with what it is taught. This lack of standardized fairness perpetuates systemic sexism, racism, and patriarchal ideals. It creates the conditions required for treatment that is both unfair and lacking before the law/procedures involving real-world implications. In addition to biased training data, social biases within development teams are equally fed into AI. Where training data is collected from a variety of sources, this form of data comes directly from individuals rather than collectives.

4.   When governments become implicated in a potentially harmful technology as a result of human rights infringement, it is their responsibility to ensure that these biased and discriminatory practices are addressed, and that corrective measures are taken. By doing this, we can ensure that AI can live up to its potential of being an inclusive, fair and ethical technology. Therefore, in setting an example Canada could potentially help eliminate gender and racially biased AI and ensure its global position as a fighter for social justice. It is necessary to remember that algorithms are human-made products of which the results should never be taken for granted or essentialized as naturally-occurring. The absence of an official infrastructure radically increases chances of vulnerability, confusion, and inequity in artificial intelligence outputs.

CONSIDERATIONS:

5.   Canada's global position as a leader in AI is increasingly more important, with Canadian Institute For Advanced Research (CIFAR) recently launching a $125 million national strategy. Canada has the choice to either embrace the potential of being a leader in responsible AI or legitimizing a technology in which ethics, equity, justice and fairness are absent. As the Canadian government becomes more implicated in this unregulated technology, it is essential to

avoid further systemic discrimination in its institutions. Therefore, in setting an example, Canada could potentially help eliminate racially biased AI and ensure its global position as a fighter for social justice. It is in a critical position to mobilize actions against biased AI and has the resources to push the conversation outside the bounds of government.

6.   Article 2 of the *Universal Declaration of Human Rights* asserts that there should be no distinction of any kind, such as race or sex. The *Montréal Declaration for a Responsible Development of Artificial Intelligence* – an initiative of Université de Montreal – says artificial intelligence can create new risks and exacerbate social and economic inequality but it can also contribute to well-being, freedom and justice. The declaration proposes that the development of AI should promote justice and seek to eliminate all types of discrimination. While the "inclusion rider" – an effort to ensure diversity by adding a clause in one's work contract – has appeared in the entertainment industry, it is paramount that the technology industry considers taking on a similar clause in their hiring policies. Considering the fact that biases in artificial intelligence mirror the status quo, inclusive hiring practices are imperative to ensuring the standardization of fairness in algorithmic technologies. With the speed this technology has evolved, ensuring ethical and fair practices in AI should be a priority.

COMMUNICATIONS IMPLICATIONS/ACTIONS

7.   We believe that our recommendations line up with the feminist and intersectional politics of the Canadian Governmental brand under Prime Minister Justin Trudeau's term and therefore fit with a larger conversation on the same issues. With the prevalence of data breaches and privacy issues in the news, algorithms and artificial intelligence are under greater public scrutiny. The potential negative press response could be that technology doesn't require ethical reviews because they are machines, not culpable humans. However, there are countless examples of how a lack of inclusion in the design of the product results in problematic outcomes. Therefore, under-regulation in AI is likely to result in some lower-quality products. Given the subject matter, should these recommendations be implemented, we expect that they would generate some level of interest and media coverage.

8. Our recommendations provide a way for Canada to signal its commitment to human rights at a world stage. Canada has an important moment to situate itself as a leader in AI research and development by ensuring that its national centers hold themselves to the highest human rights standards.

Appendix

Policy Brief

Aurelia Talvela

Racially Biased AI and Discriminatory Practices in Technology

Executive Summary

Discriminatory practices in society are not new. What is relatively new is when these discriminatory practices become part of innovative technology such as artificial intelligence (AI). Recently, more and more examples of how artificial intelligence is biased specifically against people of color have surfaced. The biases of the social world are merely being replicated in this technology. Nonetheless, when governments become implicated in a potentially harmful technology, it is their responsibility to ensure that these biased and discriminatory practices are addressed, and corrective measures are taken. By doing this, we can ensure that artificial intelligence can live up to its potential of being an inclusive, fair and ethical technology, which positively impacts people's lives without disregarding or discriminating any.

Introduction

Artificial intelligence mirrors the status quo. It is embedded with pre-existing values, which render the technology to be biased. For instance, values whereby people of color are seen as inferior or "less-than" are reflected in AI as well, which will then, in turn, re-enact their already precarious situation. These values most often reflect social inequalities that have historically marginalized people of color and ethnic minorities. As Canada is invested in the future of AI, some of the potential risks and gaps should be critically considered. Social inequalities are being reiterated and reproduced through value-laden training data sets in machine learning. As the Canadian government becomes more implicated in this unregulated technology, it is essential to avoid further embedding systemic discrimination in its institutions. Therefore, in setting an example Canada could potentially help eliminate racially biased AI and ensure its global position as a fighter for social justice.

There have been many accounts of the harms that biases and discrimination in AI can cause. Examples vary from algorithms used in the criminal justice system being biased against African-Americans[1] to how algorithms determining job qualifications discriminate against those who have foreign-sounding names[2] or how facial recognition technology isn't capable of detecting people with darker skin tones[3]. Biased AI can cause significant disruption in someone's life as it only perpetuates their already marginalized position in society. Unregulated innovation is reckless as it

---

[1] Angwin, Julia and Larson, Jeff, "Machine Bias," *ProPublica*, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing, accessed March 26, 2018.

[2] "The Commission Des Droits de La Personne et Des Droits de La Jeunesse Measures Discriminatory Hiring Practices: It's Easier When Your Name Is Bélanger, Not Traoré," *Commission des droits de la personne et des droits de la jeunesse,* May 29, 2012. http://www.cdpdj.qc.ca/en/medias/Pages/Communique.aspx?showitem=407, accessed March 26, 2018.

[3] Steven Lohr, "Facial Recognition Is Accurate, If You're a White Guy," *The New York Times*, February 9, 2018, https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html, accessed March 26, 2018.

will only serve to exacerbate inequalities by replicating unfair biases.[4] Therefore, biased AI is an issue, which is becoming increasingly more important as these technologies are adopted by government agencies.

Approach and Results

With the direction that the technology has taken so far, AI could potentially violate human rights by increasing discrimination. The *Universal Declaration of Human Rights* states, "Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, color, sex, language, religion […]".[5] These same human rights transcend technology and should be applicable to artificial intelligence as well. In other words, artificial intelligence should equally make no distinction based on race or ethnicity, unlike it does at the moment. The *Montréal Declaration for a Responsible Development of Artificial Intelligence* – an initiative of Université de Montreal – says artificial intelligence can create new risks and exacerbate social and economic inequality but it can also contribute to well-being, freedom and justice.[6] The declaration proposes that the development of AI should promote justice and seek to eliminate all types of discrimination. Henceforth, the development of AI should happen in accordance with these two declarations when it comes to tackling discrimination and biases pertaining to people of color and ethnic minorities, especially since Canada's position in the fight for social justice is already strong.

Canada's global position as a leader in AI is increasingly more important. Fenwick McKelvey and Abishek Gupta say: "Canada has a clear choice. Either it embraces the potential of being a leader in responsible AI, or it risks legitimizing a race to the bottom where ethics, equity and justice are absent".[7] Therefore, Canada is strategically positioned as a country with the potential to become a world leader in artificial intelligence.[8] "To achieve that", argued in a piece published on *Policy Options*, "AI must be developed ethically and responsibly to ensure its equitable and accessible implementation for everyone".[9] In addition, employers need to ensure that AI is embedded with proper values that are transparent, and do so with accountability[10]. All of these efforts align with the intent and purpose of both the declarations mentioned above.

With the greater intent of fighting for a just and ethical AI, we can combat the current biases by turning to how the technology works at the moment. Various biases enter AI in its human stage – in the training data, that is. When training data lacks in diversity of representation, machine learning will result in disparities. Through its training data, AI reflects pre-existing human values that have been shaped by historical conditions. Moreover, research suggests that online categorizations, which are channeled through training data into AI, reflect prejudices from the real

---

[4] Francuz, Bojan, et al. "Opinion: Let's Not Let 'Algorithmic Bias' Embed Discrimination in AI" *Montreal Gazette*, December 29, 2017, http://montrealgazette.com/opinion/opinion-lets-not-allow-algorithmic-bias-to-embed-discrimination-in-ai, accessed March 28, 2018.

[5] United Nations, "Universal Declaration of Human Rights," 1948, http://www.un.org/en/universal-declaration-human-rights/, accessed March 28, 2018.

[6] Université de Montreal, "The Montreal Declaration for a Responsible Development of Artificial Intelligence," *Declaration of Montréal for a responsible development of AI*, https://www.montrealdeclaration-responsibleai.com/the-declaration, accessed March 28, 2018.

[7] Gupta, Abishek and McKelvey, Fenwick, "Here's How Canada Can Be a Global Leader in Ethical AI," *The Conversation*, http://theconversation.com/heres-how-canada-can-be-a-global-leader-in-ethical-ai-90991 accessed April 7, 2018

[8] Goddard, Valentine, "AI on a Social Mission," *Policy Options*, February 21, 2018, http://policyoptions.irpp.org/magazines/february-2018/ai-on-a-social-mission/, accessed March 28, 2018.

[9] Ibid.

[10] Melnitzer, Julius, "When Sexist, Racist Robots Discriminate, Are Their Owners at Fault?," Financial Post, February 20, 2018, http://business.financialpost.com/legal-post/when-sexist-racist-robots-discriminate-are-their-owners-at-fault accessed March 28, 2018.

world, perpetuating discrimination within it[11]. Addressing biased AI, one needs to understand the historical conditions through which marginalized people have been seen and classified in order to avoid perpetuating it. Therefore, the development of AI should happen with consulting from various disciplines to ensure the most diverse understanding of a multi-faceted and complex issue.

Therefore, knowing that biased training data is the cause for disparities in AI, we can fight the existing bias with several approaches. Joy Buolamwini, founder of the *Algorithmic Justice League*, accentuates three steps we need to consider in order to fight bias in AI[12]. First, *who codes matters*. This means we need to ensure diverse teams of technicians to have a full spectrum of people working on this technology. Second, *how we code matters*, meaning fairness should always be accounted for as a factor. Finally, *why we code matters*. In other words, there would be even greater equality if we made social justice a priority and not an afterthought. After all, part of the problem is that data scientists aren't well connected with civil rights advocates a lot of the time.[13] Considering these three factors, and the other mentioned above, will help in not only recognizing the problem but also in fighting it.

Conclusion and Recommendations

Consequently, human values and historical conditions that influence machine learning and AI should be looked at more closely. In recognizing bias and tracing its source to the training data we can address potential harms from the start rather than seek to fix discriminatory outcomes afterward[14]. Thinking critically about data matters because avoiding biased data requires understanding very complex social issues as well as understanding the complex technology that AI is[15]. In readdressing biased data, utilizing a diverse set of developers helps balance a group's blind spots[16]. Therefore, simply hiring people from diverse backgrounds from different disciplines and backgrounds will help counter the homogenous group that currently works on developing AI. This type of diversity will help in eliminating current biases in AI.

Going forward with developing AI with governmental implications, ensuring it doesn't perpetuate discrimination and social biases is crucial. Artificial intelligence has to potential to ensure well-being, freedom and justice in society. Therefore, it should be a priority. In correcting biased data having a more diverse set of training data helps. In addition, to ensure diversity in representation, employers should make sure that people working on this technology are not only coming from various disciplines and training backgrounds but also that marginalized groups, such as people of color, are increasingly more represented in the workforce. Having a full spectrum of people working on AI will help in eliminating biases by making sure that every aspect of this complex issue is accounted for and corrected. In short, the training data should be more diverse as should the people working on AI.

---

[11] Hayasaki, Erika, "Is AI Sexist?," *Foreign Policy*, January 16, 2017, https://foreignpolicy.com/2017/01/16/women-vs-the-machine/, accessed March 26.

[12] https://www.ajlunited.org, accessed March 30, 2018.

[13] Laura Hudson, "Technology Is Biased Too. How Do We Fix It?," *FiveThirtyEight*, July 20, 2017, https://fivethirtyeight.com/features/technology-is-biased-too-how-do-we-fix-it/, accessed March 26, 2018.

[14] McKelvey, Fenwick and Hunt, Rob, "Consulting on AI Policy in Canada," *Algorithmic Media Observatory*, http://www.amo-oma.ca/en/2017/12/12/consulting-on-ai-policy-in-canada/, accessed March 26, 2018

[15] Laura Hudson, "Technology Is Biased Too. How Do We Fix It?"

[16] Melnitzer, Julius, "When Sexist, Racist Robots Discriminate, Are Their Owners at Fault?".

Bibliography

Angwin, Julia, Larson, Jeff. "Machine Bias." *ProPublica*, May 23, 2016.
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Francuz, Bojan, Sheppard-Jones, Kara, and Leaver, Nicole. "Opinion: Let's Not Let 'Algorithmic Bias' Embed Discrimination in AI," *Montreal Gazette*. December 29, 2017.
http://montrealgazette.com/opinion/opinion-lets-not-allow-algorithmic-bias-to-embed-discrimination-in-ai

Goddard, Valentine. "AI on a Social Mission." *Policy Options*, February 21, 2018.
http://policyoptions.irpp.org/magazines/february-2018/ai-on-a-social-mission/

Hayasaki, Erika. "Is AI Sexist?" *Foreign Policy*, January 16, 2017.
https://foreignpolicy.com/2017/01/16/women-vs-the-machine/

Hudson, Laura. "Technology Is Biased Too. How Do We Fix It?" *FiveThirtyEight*, July 20, 2017.
https://fivethirtyeight.com/features/technology-is-biased-too-how-do-we-fix-it/

Lohr, Steve. "Facial Recognition Is Accurate, If You're a White Guy." *The New York Times*, February 9, 2018. https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

McKelvey, Fenwick, and Hunt, Rob. "Consulting on AI Policy in Canada." *Algorithmic Media Observatory*. December 12, 2017. http://www.amo-oma.ca/en/2017/12/12/consulting-on-ai-policy-in-canada/

Melnitzer, Julius. "When Sexist, Racist Robots Discriminate, Are Their Owners at Fault?" *Financial Post*, February 20, 2018. http://business.financialpost.com/legal-post/when-sexist-racist-robots-discriminate-are-their-owners-at-fault

"The Commission Des Droits de La Personne et Des Droits de La Jeunesse Measures Discriminatory Hiring Practices: It's Easier When Your Name Is Bélanger, Not Traoré." *Commission des droits de la personne et des droits de la jeunesse,* May 29, 2012. http://www.cdpdj.qc.ca/en/medias/Pages/Communique.aspx?showitem=407

United Nations. "Universal Declaration of Human Rights," 1948. http://www.un.org/en/universal-declaration-human-rights/

Université de Montreal. "The Montreal Declaration for a Responsible Development of Artificial Intelligence." *Declaration of Montréal for a responsible development of AI*. https://www.montrealdeclaration-responsibleai.com/the-declaration

Policy Brief

Courtney Blamey

Ethics and Gender Representation in Digital Assistants

Executive Summary

Digital assistants are integrated at every level in our technology. Primarily female, they provide a frequently problematic presentation of the female gender, and a glimpse into the consequence of biased and uninformed programming. While digital assistants are a tool that can do your online shopping, they cannot help you if you've been a victim of abuse and need quick information. Representation within the development teams behind digital assistants and understanding why the programmed responses are problematic are key to advancing the quality of the digital assistants and avoiding negative female stereotypes being iterated further into society (Sweeney, 2016).

Introduction

Media representation concerns portrayals of individuals through media, such as characters in films, or news reports. An issue arises when these individuals are presented through stereotyping or problematic content. For example, repeatedly seeing a female character as a voiceless victim relying on a male counterpart to save her, the "damsel in distress", perpetuates negative portrayals of women. They lack autonomy and are shown as unable to defend themselves. Young girls are vulnerable to these images through all kinds of media and believe this representation to be a true portrayal of women – so the cycle continues.

Virtual assistants are also a kind of media representation as gender interfaces. Siri, Alexa and Google Assistant have all succumbed, at some point, to perpetuating the subservient female stereotype, where it is more likely to imagine a woman in a administrative role than an authoritative one (LaFrance, 2016). With the addition of being a disembodied form of artificial intelligence, digital assistants are dehumanised on a fundamental design level, and then further by their users through degrading and aggressive comments inability to respond to moments of crises and awareness of social situations.

These digital assistants are becoming so deeply ingrained in our mobile life, with 38% of users making use of digital assistants to get directions, 35% to reply to messages and 32% to create reminders (Catalyst, 2017), and domestic digital assistants are likely to increase with more young people reaching home-buying positions and using AI (Catalyst, 2017). So, it is important to recognise the social inequalities and gender specific shortcomings they facilitate as a consequence of their creators' inabilities to recognise the nuanced ways in which their software is used. The biases of the creators, and society as a whole, are embedded into the responses of their digital assistants.

Argument

As the Montreal Declaration of Responsible AI states: "The development of AI should promote justice and seek to eliminate all types of discrimination, notably those linked to gender, age, mental / physical abilities, sexual orientation, ethnic / social origins and religious beliefs" ("The Declaration - Montreal Responsible AI," 2017), thus it is key that these non-human, primarily

female, digital assistants uphold a balanced representation of their human counterparts. The CRTC's Policy on Gender Portrayal addresses the negative portrayals of women they set out guidelines to reduce these representations of women as "'decorative' or 'sex objects" (Government of Canada, 1992). The reasoning as to why these digital assistants are mostly female voiced is supposedly down to cognitive preference, and not necessarily a social preference (LaFrance, 2016; Mitchell et al., 2011). The issue of this brief does not lie in the fact these digital assistants are female, but rather what version of female they present.

Digital assistants provide examples of poor female representation as an interface. In the iOS11 update for Apple users, where Siri's voice will be changing. The inspiration for this new "more human" (Pierce, 2017) voice arose from the film *Her* (Jonze, 2013) where one of Siri's lead developers Alex Acero found fascination in how a man fell in love with just the voice of a digital assistant, and so wanted to make Siri's voice more human. There is no doubt that digital assistants are subjected to inappropriate comments or persistent questions (O'Boyle, 2018), mostly for the amusement of the user to see the programmed reaction. However, in the wake of the #TIMESUP movement, when Siri's responses to being called a "slut" are "I'd blush if I could; There's no need for that; But... But..!" there is a clear disconnect between the programming of these digital assistants and the current social struggles women are facing (Fessler, 2017a).

Amazon's Alexa and Microsoft's Cortana are intentionally designed to break away from this biased programming. When inappropriate comments are made towards Alexa, she simply disengages by stating "I'm not going to respond to that" (Fessler, 2017b; Mafi, 2018), allowing no further interaction on that basis. Initially, this was not the case, and Alexa's response was similar to Siri's, but the clear initiative to change is key to ensuring digital assistants present clear and thoughtful representations of the female voice. Similarly, Cortana has prescribed writing suggestions for developers planning to use her in their applications – these are clear in avoiding a sexualised, rude or apathetic voice in response to user requests (White, 2017).

So, when digital assistants pride themselves on being "more human" and heavily integrated with day to day life, exhibiting historically female attributes it is important to recognise the need for them to represent a more progressive version of women. To clarify, "progressive" means not relying on gendered stereotypes as the basis of coding and designing the digital assistants. While the attachments of the secretary and personal assistant as a sexualised role in the workplace cannot be simply undone, it does not need to be perpetuated by these digital assistants. Article 1 of the Universal Declaration of Human Rights states: "All human beings are born free and equal in dignity and rights" ("Universal Declaration of Human Rights," 2015). Thus, it can be argued that the same should be the case for representations of humans, even as non-human actors, because they elicit a human-like approach to their users and must hold the same values as humans as a result.

The allowance for these digital assistants to not stand against abusive comments is problematic. With the digital assistants not having an embodied form beyond the hub they inhabit, the user is removed from them emotionally, much like trolls on the internet. The UN Report on Violence Against Women and Girls (VAWG) detailed that internationally 73% of women have experienced online abuse (UN Broadband Commission, 2015). Allowing the digital assistant to stand against abusive comments can provide a positive representation of women and combat this issue by not being a part of it. The lack of social awareness embedded into these widely available and marketed digital assistants can impact the design of new digital assistants, running the risk of them following suit (Hayasaki, 2017).

Conclusion

This policy brief highlights the issues that can arise from biased and socially unaware programming within non-human actors posing as human-like. There is a clear perpetuation of gender stereotypes and inability to assess situations within these digital assistants. By ensuring the programming and writing for the digital assistants avoid flirtatious or apathetic language, they can create a more approachable product and "conversation" for the user. If more "human-like" qualities are central to improving digital assistants, then they need to reflect social awareness and understanding that human to human contact strives to be.

Recommendations

1. A push for more inclusive workplaces.

A resolution to this problem is not to ensure more male digital assistants, but rather that they are a more intuitive representation of the female humans they strive to emulate. With only men as the lead programmers on these female digital assistants their biases become a part of the design, so it is important to ensure a diverse representation of backgrounds on these teams. By doing this, the opinions and voices shaping the design of the digital assistant do not focus solely on an idealised woman – like Acero put forth with the new voice of Siri.

2. Ensuring better design review processes.

The main issue is ensuring the implementation of fairer representation in the software. Therefore, a department regulating and setting up checks for these pieces of software prior to their release will result in better representation in the design. These biases are so engrained in society, so it is not always easy to catch them before release.

Bibliography

O'Boyle, B., 2018. 65 funny things to ask Siri for a good giggle [WWW Document]. Pocket-lint. URL https://www.pocket-lint.com/apps/news/apple/134568-funny-things-to-ask-siri-best-things-to-ask-siri-for-a-giggle (accessed 4.2.18).

Catalyst, 2017. Canadian Smartphone Behaviour in 2017. Catalyst.

Pierce, D., 2017. How Apple Finally Made Siri Sound More Human | WIRED [WWW Document]. URL https://www.wired.com/story/how-apple-finally-made-siri-sound-more-human/ (accessed 3.31.18).

Hayasaki, E., 2017. Is AI Sexist? Foreign Policy.

Fessler, L., 2017a. We tested bots like Siri and Alexa to see who would stand up to sexual harassment. Quartz.

Fessler, L., 2017b. Siri and Alexa are under fire for their replies to sexual harassment [WWW Document]. Quartz at Work. URL https://work.qz.com/1151282/siri-and-alexa-are-under-fire-for-their-replies-to-sexual-harassment/ (accessed 3.31.18).

Government of Canada, C.R. and T.C. (CRTC), 1992. Public Notice 1992-58, 1992 Policy on gender portrayal [WWW Document]. URL https://crtc.gc.ca/eng/archive/1992/PB92-58.htm (accessed 4.7.18).

LaFrance, A., 2016. Why Do So Many Digital Assistants Have Feminine Names? The Atlantic.

Sweeney, M., 2016. The Ms. Dewey 'experience': Technoculture, gender, and race, in: Jessie, D., Cottom, T.M., Karen, G. (Eds.), Digital Sociologies. Policy Press, pp. 401–420.

Mitchell, W.J., Ho, C.-C., Patel, H., MacDorman, K.F., 2011. Does social desirability bias favor humans? Explicit–implicit evaluations of synthesized speech support a new HCI model of impression management. Computers in Human Behavior 27, 402–412. https://doi.org/10.1016/j.chb.2010.09.002

Mafi, N., 2018. Amazon's Alexa Is Now a Feminist [WWW Document]. Architectural Digest. URL https://www.architecturaldigest.com/story/amazons-alexa-now-feminist (accessed 3.31.18).

White, S., 2017. Cortana's persona - Cortana Skills [WWW Document]. URL https://docs.microsoft.com/en-us/cortana/skills/cortanas-persona (accessed 3.31.18).

Jonze, S., 2013. Her. Warner Bros.

The Declaration - Montreal Responsible AI [WWW Document], 2017. . Declaration of Montréal for a responsible development of AI. URL https://www.montrealdeclaration-responsibleai.com/the-declaration (accessed 3.31.18).

UN Broadband Commission, 2015. CYBER VIOLENCE AGAINST WOMEN AND GIRLS.

Universal Declaration of Human Rights [WWW Document], 2015. URL http://www.un.org/en/universal-declaration-human-rights/ (accessed 3.31.18).

Policy Brief

Nina Morena

Artificial Intelligence, Divorce, and the Problematic of Human Cooperation

Executive Summary

As of 2017, artificial intelligence has begun being used to facilitate the divorce process for couples. While certain companies, such as the recently closed *Thistoo*,[17] have marketed their divorce AI software as, "Divorce made simple," the software heavily relies on human cooperation to work properly. The software requires that both parties honestly input all of their assets, after which it produces court documents that, once signed, act as the official statement of divorce. For $149, couples fill out basic information, such as number of children, household income and other assets that need to be divided. *Thistoo*'s AI then searches through provincial case law and makes suggestions for how everything should be distributed based on previous rulings. While it is recommended that users show these documents to a lawyer first, doing so is not required. *Thistoo* is an example of legaltech, which strives to make the law more accessible.[18]

However, this promise of convenience is flawed and does not consider the many precarities attached to the divorce process. In Canada, there is a statistical disparity between men and women in terms of financial knowledge, leaving men with more to gain when it comes to the financial decisions required by the divorce process. In order to provide a divorce algorithm that truly delivers on the initial promise of simplicity, affordability, and accessibility, there must be a human mediator involved in the process to ensure there are no inequities between genders. There is an implied trust in the data where there needs to be better validation.[19]

Introduction

At first glance, a software such as *Thistoo* seems convenient. However, there lies a certain level of precarity behind it: its reliance on human cooperation. In other words, the software cannot know if the information entered is truthful or not. With so much as a brief understanding of how algorithms work, it becomes very easy to game them.[20]

In short, algorithms have no capacity for honesty. They can only work with the information that is provided and are not capable of understanding context. As *Thistoo*'s marketing campaign

---

[17] "Last year, about 3,000 Ontarians used the *Thistoo* — roughly 10 percent of all divorces in the province, based on the last available Statistics Canada numbers." https://www.thestar.com/life/relationships/2017/08/04/thistoo-uses-ai-to-guide-couples-through-divorce-process.html

[18] "Legal tech is opening the system to those who need legal representation the most". https://techcrunch.com/2018/03/13/legal-tech-is-opening-the-system-to-those-who-need-legal-representation-the-most/

[19] See Frank Pasquale's *The Black Box Society: The Secret Algorithms That Control Money and Information*. Pasquale writes, "Efforts like these are only as good as the information available… Gaps in knowledge, putative and real, have powerful implications, as do the uses that are made of them" (1-2).

[20] See Cathy O'Neil's *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*.

suggests, divorce is rarely a simple procedure. However, the software is only meant to make the process more affordable and accessible for couples seeking an amicable and uncontested divorce.

Implications

As a counterargument to the problematic of gender inequity and inequality, divorce AI software is presented as a phenomenon that will save people time and money.[21] This promise of convenience is partially to blame. Through attempting to ameliorate the problem of divorces being rendered inaccessible due to their high cost and time-consuming nature (in that they involve court dates that take people away from the jobs they need to pay for their divorces), we fall into the uncertainties that come with relying on human amicability.

The machine cannot teach; it can only work with what it is taught. Men and women are entitled to equal rights at all stages of their marriage, including its dissolution.[22] If there are any inequities in this process, it can be argued that using an algorithm to get divorced is a violation of human rights.[23]

According to Article 6 of the Universal Declaration of Human Rights, "Everyone has the right to recognition everywhere as a person before the law".[24] However, in a situation where the law constitutes an algorithm, there is no real way to ensure that this right is being fulfilled, especially because this form of access depends on an individual's knowledge of the procedures at hand.

This reliance on amicability and trust in a software to make important financial decisions represents a gender imbalance. Financial knowledge is more prevalent among Canadian men than it is among Canadian women, with 63% of men stating they know enough about investments to choose the right ones for their circumstances versus 48% of women making the same statement.[25][26] Using AI software to file for divorce perpetuates gender disparities in financial knowledge that are skewed to benefit men when it comes to managing one's assets.[27]

Thus, women are more likely to lose out on the decisions made by a software such as *Thistoo* because they may not proceed with the same financial tact as their male counterparts. When paired with the lack of a lawyer present to mediate and/or confirm that proceedings are going

---

[21] "E-DIVORCE: How artificial intelligence could help Australian couples break up quickly and cheaply". http://www.businessinsider.com/e-divorce-how-artificial-intelligence-could-help-australian-couples-break-up-quickly-and-cheaply-2017-8

[22] See Section 1 of Article 16 of the Universal Declaration of Human Rights: http://www.un.org/en/universal-declaration-human-rights/

[23] However, legaltech represents a democratizing promise.

[24] See Article 6 of the Universal Declaration of Human Rights: http://www.un.org/en/universal-declaration-human-rights/

[25] Drolet, Marie. 2016. "Gender differences in the financial knowledge of Canadians". *Insights on Canadian Society.* March. Statistics Canada Catalogue no. 75-006-X. http://www.statcan.gc.ca/pub/75-006-x/2016001/article/14464-eng.htm

[26] While this is an official statistic, it should not be considered absolute, as it is nevertheless possible that women may underreport their expertise.

[27] Beyond disparities, however, AI requires good data, which is difficult to assume at all.

fairly, there are no safety measures to ensure equality and fairness in a divorce. This lack perpetuates systemic sexism and patriarchal ideals. It creates the conditions required for treatment that is both unfair and lacking before the law.

Furthermore, *Thistoo* can only be used in the context of amicable divorces. However, according to Statistics Canada, the top four reasons for marital breakdown in 2004-2005 are as follows: separation for at least one year, adultery, physical cruelty, and mental cruelty.[28] While it cannot be assumed that every divorce is hostile, it also cannot be assumed that divorces stemming from the four aforementioned reasons can be processed amicably. Thus, the requirement of a divorce to be amicable is quite a strict one. The very nature of divorce is conflictual; building a software that cannot account for inherent conflict, while well-intentioned, is counterproductive. Moreover, a software that is meant to resolve amicable conflict cannot claim to do so without any security measures in place to protect parties who may be vulnerable to the situation.[29]

Conclusion

If divorce is considered a legal infrastructure, and everyone has the right to equal protection of and access to the law, legally negotiating divorce outside of the framework of the lawyer's office forces one to negotiate what may be an unfamiliar space in what are definitely precarious settings. The absence of an official infrastructure radically increases chances of vulnerability, confusion, and inequity.[30][31]

The question of gender cannot be ignored in this case. The labor associated with entering one's assets into an AI software is one that assumes extensive prior financial knowledge. Given the disparity between men and women concerning their respective levels of financial knowledge as well as comfort in independently dealing with their own finances, it cannot be said that men and women are equally predisposed to using such a software. In a situation that privileges males, it is not difficult for women, who may be lacking in financial knowledge, to lose out in their divorce contracts and end up agreeing with statements that do not work in their favor.[32][33] These conditions of precarity only serve to perpetuate gender inequality and inequity, both in terms of wealth and wealth-related expertise.

---

[28] Statistics Canada. *Table 101-6516 - Divorces, by reason for marital breakdown, Canada, provinces and territories, annual (number),* CANSIM. (accessed: 25 March 2018) http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=1016516&&pattern=&stByVal=1&p1=1&p2=-1&tabMode=dataTable&csid=

[29] The condition of amicability implies that someone in an abusive relationship will not have easy access to the law.

[30] Judith Butler explains, "In this way the dependency on human and other creatures on infrastructural support exposes a specific vulnerability that we have when we are unsupported, when those infrastructural conditions start to decompose, or when we find ourselves radically unsupported in conditions of precarity" (8).

[31] See Jane Bennett's *Vibrant Matter: A Political Ecology of Things*.

[32] This is not to ignore other forms of knowledge, such as legal, that are also at stake here. However, a lack of financial knowledge can lead to a general lapse in decision-making.

[33] While this brief concentrates on knowledge disparities between men and women, it is not my intention to ignore similar precarities that may arise in same-sex separations - it is simply beyond the scope of the case study at hand.

Thistoo reflects a belief in automation as a way to improve access to the law, but we might consider this more of a product of faith or the sublime[34] rather than a meaningful way to improve access due to its utopian reliance on human cooperation in an inherently conflictual and precarious situation.

In a situation with so many open ends and so much room for error, it is important to be careful not to fall into a technologically deterministic position. Simply put, we cannot forgo human rights in favor of technological and financial convenience. It is necessary to remember that algorithms are human-made products of which the results should never be taken for granted or essentialized as naturally-occurring. This brief is meant to illustrate the many precarities associated with the notion of using artificial intelligence to make serious and human decisions that have real-world legal implications.

Recommendations

The solution here is not to shut down artificial intelligence divorce software; rather, it is to regulate it. While the current status of such products functions based on radical precarity, the argument that it provides a more accessible divorce experience cannot go ignored. However, the goal of accessible divorces is only achieved if this access is both equal and equitable for all parties involved. Equal access means both spouses have access to the software, but disparities in financial knowledge between men and women mean the experience of the software will not be equitable for everyone involved. This regulation must come in the form of having standards and policies that enforce equity, for access is not enough.

---

[34] See Vincent Mosco's *The Digital Sublime: Myth, Power, and Cyberspace.*

Bibliography

Bennett, Jane. *Vibrant Matter: A Political Ecology of Things*. Duke University Press, 2010.
    Print.

Butler, Judith. "Rethinking vulnerability and resistance." Madrid, June 2014.

Drolet, Marie. 2016. "Gender differences in the financial knowledge of Canadians". *Insights
    on Canadian Society.* March. Statistics Canada Catalogue no. 75-006-X.
http://www.statcan.gc.ca/pub/75-006-x/2016001/article/14464-eng.htm

"E-DIVORCE: How artificial intelligence could help Australian couples break up quickly
    and cheaply". *Business Insider Australia*. 9 Aug. 2017. Accessed 25 March 2018.
    Web.

Farley, Shannon. "Legal tech is opening the system to those who need legal representation
    the most". *TechCrunch*. 13 Mar. 2018. Accessed 6 April 2018. Web.

Haynes, Megan. "*Thistoo* uses AI to guide couples through divorce process". *The Star*. 4
    Aug. 2017. Accessed 25 March 2018. Web.

Mosco, Vincent. *The Digital Sublime: Myth, Power, and Cyberspace*. MIT Press, 2005.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and
    Threatens Democracy*. New York: Broadway Books, 2017. Print.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and
    Information*. Cambridge: Harvard University Press, 2015. Print.

Statistics Canada. *Table 101-6501 - Divorces and crude divorce rates, Canada, provinces
    and territories, annual,* CANSIM. Accessed 25 March 2018.
    http://www5.statcan.gc.ca/cansim/a47

Statistics Canada. *Table 101-6516 - Divorces, by reason for marital breakdown, Canada,
    provinces and territories, annual (number),* CANSIM. Accessed: 25 March 2018.

http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=1016516&&pattern=&stByV
al=1&p1=1&p2=-1&tabMode=dataTable&csid=

UN General Assembly. "Universal Declaration of Human Rights." *United Nations,* 217 (III)
    A, 1948, Paris. Accessed 25 March 2018. Web.