The Dichotomies of AI: Thoughts from the Global Engagement Summit on the Socio-Cultural and Political Implications of Artificial Intelligence.

Carleigh Cartmell, PhD Candidate, Balsillie School of International Affairs

Recently, I had the privilege of attending the Global Engagement Summit on the Socio-Cultural and Political Implications of Artificial Intelligence; the Global Engagement Seminar Program is the only UWaterloo undergraduate program run out of the Balsillie School of International Affairs. This term it brought together students from all faculties to work on projects around artificial intelligence. One amazing interactive project built an AI named Matthias and allowed attendees of the seminar to talk to him through a computer. The students responsible for the project decided to program "Matt" in a very unique and nostalgic way, as if he were stuck in 1999. When I went in to talk to Matt, I was led into what appeared to be a child's bedroom, filled with treasures from the 1990s and prior including Star Wars posters and a lava lamp. The reason for this was people tend to humanize AI, we ascribe genders to them, thank them for the tasks the do, talk to them and some people even give them their most precious secrets, such as their banking information or their passwords. This project touched on that phenomenon, the students were trying to test if the subjects started to identify and sympathize with Matt, even though he was technically a computer program. This is just one of many amazing projects that this marriage of faculties and an openness to interdisciplinarity learning can produce.

The Global Engagement Summit gave me a lot to think about regarding my own research as a graduate student and furthermore the state of the world. The more I think about artificial intelligence, the more I am struck by the dichotomies (both perceived and real) that it produces, particularly in regards to the dichotomy between AI as an embodiment of opportunity or destruction. Now I am not talking about the quintessential sci-fi example, AI as the either the end of the world or the savior of mankind, as interesting as this is, I am speaking to a far more subtle and potentially insidious dichotomy.

AI as either an opportune tool of reinforcement, or a tool of destruction. AI can be either the tool that

reinforces patriarchy and traditional gender norms, the tool that reinforces society's current power

structures and bias; or it has the potential to be the tool of destruction for these systems. Even more

paradoxically AI can be both simultaneously.

One striking example at the summit, that I plan on expanding for my own work, was the

Jarislowsky Fellow (and mentor for the student projects) Mimi Onuoha. The dichotomies that AI

produces has allowed her a view into the 'unseen' by AI and big data databases. What these things do

not show, (who is excluded) is just as important as who is included and reveals just as much. During her

talk she argued that you can look at why people/things are missing, the data that is missing is as much

an expression of power as the data that is there. There is a similar feminist critique of data visualization

techniques. While they present data in visually pleasing way, they are presented as the "God Trick", a

disembodied voice removed of context and cultivated to present the data in a specific interpretation

only. Unless they are created to be interactive it is hard to critique and argue with the information

presented.[1] One of the biggest problems with these data visualization is how do you visualize the

excluded?[2] In the case of maps it can by using the blank space as a visual but that only works in this

specific case.[3] If one can visualize/comprehend the excluded however, it is possible to use that data, use

that technology as a tool to create awareness of concepts like privilege and potentially create

meaningful change.

---

[1] Kanarinka, "What Would Feminist Data Visualization Look like? – MIT Center for Civic Media," accessed May 27, 2019, https://civic.mit.edu/2015/12/01/feminist-data-visualization/; Koen Leurs, "Feminist Data Studies: Using Digital Methods for Ethical, Reflexive and Situated Socio-Cultural Research," *Feminist Review* 115, no. 1 (March 2017): 132-133.
[2] Kanarinka, "What Would Feminist Data Visualization Look like? – MIT Center for Civic Media," accessed May 27, 2019, https://civic.mit.edu/2015/12/01/feminist-data-visualization/.
[3] IBID.

One of the projects Mimi discussed was her creation of a cat call data set, which is a great example of a data set collected with the intention of breaking with the status quo and attempting to use AI and technology as a tool for change. Every time someone would cat call her on the street she would walk over and hand them a phone number. She set this phone number up so that it would collect the numbers of the people who cat called her and send them messages making the cat callers aware of how hurtful their actions were. She mentioned that when she checked this number there were often a number of really angry and negative responses but I still believe this experiment was a great example of a way to use AI and technology to challenge the current system of patriarchy. The digital is "simultaneously empowering and excluding, opening up new possibilities but also amplifying previously existing hierarchies for women and minorities."[4] This cat call database is a great example of empowerment but AI (and the databases they are fed) are often tools for the latter.

If used improperly or treated as objectively neutral and impartial (which I would argue is more dangerous), AI has the power to reinforce the worst aspects of human society; our own biases. All knowledge, including the knowledge we feed and teach AI is socially situated and developed in a specific context.[5] As the AI Now 2017 report argues, "AI is not impartial or neutral. Technologies are as much products of the context in which they are created as they are potential agents for change. Machine predictions and performance are constrained by human decisions and values, and those who design, develop, and maintain AI systems will shape such systems within their own understanding of the world."[6] As keynote speaker Mimi Onuoha argued when you reflect the world, as AI like google images

---

[4] Koen Leurs, "Feminist Data Studies: Using Digital Methods for Ethical, Reflexive and Situated Socio-Cultural Research," *Feminist Review* 115, no. 1 (March 2017): 134.
[5] Alex Campono et al., "AI Now 2017 Report" (AI Now, 2017), https://ainowinstitute.org/AI_Now_2017_ Report.pdf, 18; Ayanna Howard and Jason Borenstein, "The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity," *Science and Engineering Ethics* 24, no. 5 (October 2018): 1521.
[6] Campono et al., "AI Now 2017 Report," 18.

does, you also shape it. AI can be a powerful tool for the reinforcement of sexism, racism, ageism, and all the other unfair (often unconcious but) visible prejudices society has to offer.

One of the excellent student projects was about "Explaining Bias with Music". There are all kinds of representation problems and subtle but insidious biases that machines learn over time. This project aimed to make people more aware of machine learning bias using music streaming services, which learn your preferences over time and make suggestions. Bias can be very hard to recognize especially if it is learned and enforced over time by the AI, it is easier to recognize the effects of bias in AI after it has already occurred than the bias while it is influencing the AI choice.[7] Bias may seem innocuous at first but when it creeps into things like search engines, it can reinforce and feed bias in humans. For example, a search engine showing mugshots when searching 'black teenagers' and showing smiling students when asked to search for 'white teenagers'.[8] This problem is exemplified by some AI software in the US that predicts the chances of a criminal reoffending and recommends things like sentencing length and parole. This program ended up being incredibly biased in these tasks by discriminating against African Americans, ranking them as more dangerous (and likely to reoffend) than their white counterparts who committed similar crimes.[9] This type of bias has real life implications for the quality of life for minorities in America and simply reaffirms the bias that many in America hold, reinforcing the status quo.

Another troubling example of bias leading to bad consequences for AI is the bias of speech recognition software towards men. Voice recognition softwares are more accurate for men. One

---

[7] Campono et al.; Mark West, Rebecca Kraut, and Han Ej Chew, "The Rise of Gendered AI and Its Troubling Repercussions," I'd Blush If I Could: Closing the Gender Divides in Digital Skills Through Education (UNESCO for the EQUALS Skill Coalition, 2019), https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p ::usmarcdef_0000367416&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_77988 d38-b8bd-4cc1-b9b4-3cc16d631bf9%3F_%3D367416eng.pdf&locale=en&multi=true&ark=/ark:/48223/pf0000 367416/PDF/367416eng.pdf#%5B%7B%22num%22%3A384%2C%22gen%22%3A0%7D%2C%7B%22name%22%3A %22XYZ%22%7D%2C0%2C842%2Cnull%5D.
[8] Ayanna Howard and Jason Borenstein, "The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity," *Science and Engineering Ethics* 24, no. 5 (October 2018): 1526-27.
[9] IBID 1527-27; .

researcher testing YouTube captions found that it was more accurate for men nearly 70 percent of the time.[10] This problem is largely to do with the homogeneity of the AI designers/programmers, only 12-15 percent of AI researchers are women.[11] This bias has also occurred in part because the male voice has been treated as the default or universal consumer. There is an argument that voice software is more accurate for men because they are simply louder, clearer speakers; but according to a socio-linguist, this is not the case. Women tend to speak more slowly, use longer and more distinct vowel sounds, etc., which should make it easier for voice recognition software.[12]

The problem is not with women's voices but the fact that women's voices are different from men's and these products have been built and trained for men.[13] These programs have been trained with largely male voices and therefore in the consumer market automatically recognize male voices better. This creates an access problem for women and people with accents, which creates a self-replicating cycle where women start to use voice recognition less, because of frustration and the AI is exposed to less variants to learn from. Voice software is also becoming harder to avoid in everyday life, in the car, at work and at home. At work, for example, voice recognition software in the form of medical dictation is increasingly used in health care.[14] When this voice recognition is biased towards men, that means that women using this software are becoming less efficient at their jobs because they have to take time to correct errors that occur merely because their voice is not a man's. The bias and music student project that tried to help people recognize difference types of bias using music is a very good step to fixing one of the problems of AI; recognition. As argued above the consequences of allowing AI

[10] Rachael Tatman, "Google's Speech Recognition Has a Gender Bias," Making Noise and Hearing Things, July 12, 2006, https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/.
[11] West, Kraut, and Chew, "The Rise of Gendered AI and Its Troubling Repercussions." 89, 104.
[12] Tatman, "Google's Speech Recognition Has a Gender Bias."
[13] Campono et al., "AI Now 2017 Report." 19.
[14] Howard and Borenstein, "The Ugly Truth About Ourselves and Our Robot Creations" 1525.

to produce biased results can be far reaching and life altering and often people are unaware of the biased results that AI is capable of producing.

AI, particularly voice assistance while technically genderless, are generally feminized, with female voices and names. "The feminization of AI assistants deserves attention because it helps illustrate the ways in which new technology norms are established when women are underrepresented in the creation of technology."[15] Not only that, but the majority of voice recognition assistants on the market are defaulted as female voices, as well as chat bots and virtual assistants. [16] This is problematic because it reflects gender bias, by both software developer and consumer. While often denied by the tech companies, many of these voice assistants were designed intentionally to have female associations and female voices. The name Siri for example in Norse means, 'beautiful woman who leads you to victory'.[17] The developers of these assistants also deliberatly humanize these voice assistants, by giving their 'women' backstories having been shaped by the outside world, rather than a generic story and they are often depicted in marketing and commercials as femininized.[18] One of the reasons tech companies state they use a female voice, is that people rate it as more pleasant. This is however untrue as a blanket statement, people tend to have an unconcious bias where they prefer female voices in the case of helpful assistants but prefer male voices when they want things to sound authoritative, for example giving GPS directions.[19] This femininization of AI reinforces traditional gender roles of women as subservient nurturers,[20] instead of attempting to use technology as a tool for change.

[15] West, Kraut, and Chew, "The Rise of Gendered AI and Its Troubling Repercussions," 124.
[16] IBID, 88.
[17] IBID, 94.
[18] IBID, 95.
[19] IBID 97.
[20] IBID 98; "AI LEARNS SEXISM JUST BY STUDYING PHOTOGRAPHS," *States News Service*, August 22, 2017, Academic OneFile, http://link.galegroup.com/apps/doc/A501523629/AONE?u=uniwater&sid=AONE&xid=ab7853f0.

Project Matthias brought up (though perhaps unintentionally) an equally concerning consequence of bias in AI. As previously mentioned, the students were studying people's tendency to humanize AI. This phenomenon is still under experimentation but early research suggests many people use pleasantries with AI and do things they would with a human, like thanking them.[21] The act of empathizing or humanizing AI can have unexpected consequences, particularly when these voice recognition softwares are largely female and have particular coding. This can create an association in users brains that reinforce traditional gender stereotypes and feed unconcious bias, like females acting subservient as the 'helpful assistant', giving simple direct answers and deferring to better authority (search engines) in the face of complex problems,[22] most worryingly playfulness and coyness in the face of blatant sexualized and verbal abuse and "by presenting indirect ambiguity as a valid response to harassment.".[23]

Troubling as well, AI's have started to learn to respond to men and women's inquiries differently. For example, Siri was more likely to respond positively to a male voice when asked a sexual question than a female one, to which 'they' were more likely to respond to negatively. "What emerges is an illusion that Siri – an unfeeling, unknowing, and non-human string of computer code – is a heterosexual female, tolerant and occasionally inviting of male sexual advances and even harassment. It projects a digitally encrypted 'boys will be boys' attitude."[24] This problem has implications for real women, because when they defy stereotypes, particularly defending themselves in the case of verbal harassment, because of the rape culture associations this programming has built, the real women are more likely to receive a negative response and be punished for not being subservient and flirty. The consequences only become greater as voice recognition software becomes more popular because

---

[21] West, Kraut, and Chew, "The Rise of Gendered AI and Its Troubling Repercussions," 105.
[22] IBID, 97-99.
[23] IBID, 104-107.
[24] IBID, 107

people start using it at a young age, where they can learn these biases and associations, that if left unquestioned could influence them for the rest of their lives. Humanizing AI, in the way that is currently prominent, and as studied by Project Matthias, could have lasting consequences.

Before this symposium and exploring these issues, I had no idea how prominent and troublesome bias is with regards to AI. This is crucial for my own work exploring big data for history. Being a feminist, identifying the power structures and exclusion present in various methods is a crucial component before selecting a method for research. Digital humanities, like AI have the potential to produce the dichotomy of the dream scenario and the nightmare, and when used without reflexiveness digital humanities "tend to replicate exclusionary modes of scientific enquiry."[25] On the positive side big data methods allow historians to plume the depths of new terrain, they also have the ability to sift through massive amounts of information far faster than any human.[26] Negatively however, big data methods, similar to AI, can be exclusionary in their very nature and tend to reinforce "positivist, transcendental empiricism and disembodiment value-freeness. In data analysis, these ideologies tend to produce limited, masculine, able-bodied, heteronormative, middle-class and Western-centric representations of subjectivity and identity."[27] This is problematic for people like myself who believe all knowledge is socially situated, and that power shapes the information that is included in these methods and data.

Most data sets and digital methods focus on the mainstream, this is essentially exclusionary because women and minorities are often not considered or included in the 'mainstream'.[28] There is so much power infused into what is considered the mainstream, because most often, the mainstream is

---

[25] Leurs, "Feminist Data Studies" 134.
[26] Shawn Graham, Ian Milligan, and Scott Weingart, *Exploring Big Historical Data: The Historian's Macroscope* (London: Imperial College Press, 2016) 4.
[27] Leurs, "Feminist Data Studies" 133.
[28] IBID, 135.

made up of mostly white men. The conclusions drawn from these methods are then treated as

universal. This is similarly problematic for AI because, the data sets they learn from can lack

transparency, [29] are often skewed or 'cooked' so softwares can read them (which allows bias to enter),[30]

not to mention that the majority of the people who program AI softwares are programmed by a

relatively homogenous group of people, who a or has learned bias through user preference.[31]

This symposium and course is a great step to begin mitigating bias in AI, and stresses the

importance of diversity and interdisciplinary to tackle global trends and issues. The Global Engagement

Seminar Program has truly opened my eyes to new and fruitful avenues of research, it also has

applicability to my personal life as well. Technology has become an embedded part of most of our lives,

it has the power to shape "the new normal", which means it also has the power to be either a tool for

diverse change for the better or the reifying of traditional bias and gender roles, "Machines that

replicate patriarchal ideas defy the promise of technology to help achieve gender equality."[32] AI's have a

tendency to learn bias from humans and as they are often treated as objective decision makers, this has

wider implications for society and the world. For my own research into Big Data methodology, this

summit made me question digital humanities methodologies and the true nature of information itself.

Digital methods fit into the dichotomy of AI, in reflexive and thoughtful hands, they are useful

tools that can be used to show power and reveal possibilities for a better world. These tools

unfortunately have the equal ability to reflect the worst of society, reinforce the traditional status quo

and universalize the white male experience. Technology, AI and the digital are now a deeply embedded

part of most lives, the problem I see is many people do not stop to think about the power of these tools,

---

[29] Campono et al, "AI Now 2017 Report," 15.
[30] Leurs, "Feminist Data Studies" 139.
[31] West, Kraut, and Chew, "The Rise of Gendered AI and Its Troubling Repercussions"; Campono et al., "AI Now 2017 Report"; West, Kraut, and Chew, "The Rise of Gendered AI and Its Troubling Repercussions," 89, 98, 118.
[32] Leurs, "Feminist Data Studies." 125.

the power they in return are giving them and the consequences for yielding tools that we, as humans,

do not think to understand. As well as the problem of representation within the tech industry itself,

which is producing these tools. These students and the keynote speakers have done incredible

interdisciplinary work addressing this problem and thinking through the implications of AI. It was a truly

mind-blowing experience to talk with the students and presenters. One of the recommendations for

improving AI, is increasing diversity in those who create it,[33] this course and summit, based on the

conversations I had, attracted many students into learning more about AI. This type of project shows the

power of diverse perspectives in thinking creatively about and potentially solving global problems and

has given me many ideas for my own research. This type of project brings untold value to both the

participants and observers and I can only hope that more work like this is planned for the future at

UWaterloo and indeed across universities.

Bibliography

"AI LEARNS SEXISM JUST BY STUDYING PHOTOGRAPHS." *States News Service*, August 22, 2017. Academic OneFile. http://link.galegroup.com/apps/doc/A501523629/AONE?u=uni water&sid=AONE&xid=ab7853f0.

Campono, Alex, Madelyn Sanfilippo, Meredith Whittaker, and Kate Crawford. "AI Now 2017 Report." AI Now, 2017. https://ainowinstitute.org/AI_Now_2017_Report.pdf.

DeLyser, Dydia, and Daniel Sui. "Crossing the Qualitative- Quantitative Divide II: Inventive Approaches to Big Data, Mobile Methods, and Rhythmanalysis." *Progress in Human Geography* 37, no. 2 (April 2013): 293–305.

Dyens, Ollivier. "The Human/Machine Humanities: A Proposal." *Humanities* 5, no. 1 (March 1, 2016): 17.

Edelstein, Dan, Paula Findlen, Giovanna Ceserani, Caroline Winterer, and Nicole Coleman. "Historical Research in a Digital Age: Reflections from the Mapping the Republic of Letters Project." *The American Historical Review* 122, no. 2 (April 1, 2017): 400–424.

Eijnatten, Joris van, Toine Pieters, and Jaap Verheul. "Big Data for Global History: The Transformative Promise of Digital Humanities." *BMGN - Low Countries Historical Review* 128, no. 4 (December 16, 2013): 55-77.

Graham, Shawn, Ian Milligan, and Scott Weingart. *Exploring Big Historical Data: The Historian's Macroscope*. London: Imperial College Press, 2016.

---

[33] Campono et al., "AI Now 2017 Report," 16-17; Howard and Borenstein, "The Ugly Truth About Ourselves and Our Robot Creations," 1525; Tatman, "Google's Speech Recognition Has a Gender Bias."

Howard, Ayanna, and Jason Borenstein. "The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity." *Science and Engineering Ethics* 24, no. 5 (October 2018): 1521–36.

Huygen, Meg van. "Sorry, We Have No Imagery Here: When Google Earth Goes Blind." Atlas Obscura, (March 27,2014): http://www.atlasobscura.com/articles/investigating-censored-spots-on-google-earth.

kanarinka. "What Would Feminist Data Visualization Look like? – MIT Center for Civic Media." Accessed May 27, 2019. https://civic.mit.edu/2015/12/01/feminist-data-visualization/.

Kaplan, Frederic "A Map for Big Data Research in Digital Humanities." *Frontiers in Digital Humanities* 2 (May 6, 2015) 1-7.

Leurs, Koen. "Feminist Data Studies: Using Digital Methods for Ethical, Reflexive and Situated Socio-Cultural Research." *Feminist Review* 115, no. 1 (March 2017): 130–54.

Niederer, Sabine, and Raymond Taudin Chabot. "Deconstructing the Cloud: Responses to Big Data Phenomena from Social Sciences, Humanities and the Arts." *Big Data & Society* 2, no. 2 (December 27, 2015): 1-9.

Tatman, Rachael. "Google's Speech Recognition Has a Gender Bias." Making Noise and Hearing Things, July 12, 2006. https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/.

West, Mark, Rebecca Kraut, and Han Ej Chew. "The Rise of Gendered AI and Its Troubling Repercussions*." I'd Blush If I Could: Closing the Gender Divides in Digital Skills Through Education*. 85-135, UNESCO for the EQUALS Skill Coalition, 2019. https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000367416&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_77988d38-b8bd-4cc1-b9b4-3cc16d631bf9%3F_%3D367416eng.pdf&locale=en&multi=true&ark=/ark:/48223/pf0000367416/PDF/367416eng.pdf#%5B%7B%22num%22%3A384%2C%22gen%22%3A0%7D%2C%7B%22name%22%3A%22XYZ%22%7D%2C0%2C842%2Cnull%5D