AI's Learned Biases Perpetuate Gender Bias and Heteronormativity

While artificial intelligence continues to drive humanity forward, the algorithms used to create it remain rooted in the past. Programmers who build AI instill their own biases into the machine learning algorithms. Their efforts result in an AI that projects and reinforces learned biases to its audience. Facial recognition technology to identify gender and conversational chatbots like Mitsuku demonstrate gender bias in AI to serve as reminders of the heteronormative, gender binary world in which we live.

Facial recognition software often ignores the nuanced web that makes up gender and instead opts for a simpler binary approach. Computer language's binary structure lends itself to categorize gender binarily, too. Humans associate certain characteristics with certain genders. For example, men have facial hair and women wear makeup.[1] Facial recognition AI inherits these same associations and struggles to think outside of the gender binary box. In a study published earlier this year, Computer science researchers at the University of Colorado, Boulder, found that AI based facial analysis tools routinely fail to identify trans and non-binary people.[2] The facial analysis tool used in the study correctly identified cisgender men and women 98% of the time. But the tool fell flat outside of the cis realm only identifying 70% of trans people correctly. The percentages dropped even farther for non-binary, agender, and genderqueer people. The tool correctly identified 0% of non-binary, agender, and genderqueer people. Some might say that the AI's prioritization of the common cisgender case falls in line with the saying "Make the common case fast," a popular saying in the tech world. But by making the common

---

[1] Dozier, Raine. "Beards, Breasts, and Bodies." *Gender & Society* 19, no. 3 (2005): 297–316. https://doi.org/10.1177/0891243204272153.

[2] Milton, Josh, and Vic Parsons. "Facial Recognition Software Can't Identify Trans People, According to Science." PinkNews, October 21, 2019. https://www.pinknews.co.uk/2019/10/21/facial-recognition-software-cant-identify-trans-people-science/.

case right, and the uncommon case wrong, facial recognition AI further marginalizes gender minorities. Given facial recognition's growing presence within society, AI's gender bias has the legs to run rampant.

Facial recognition continues to gain traction in law enforcement, immigration services, banks, social media, and other institutions, and has the potential to oppress people who do not look like 'typical' men and women. Facial recognition has made strides toward moving into gendered spaces like bathrooms and locker rooms. China already has facial recognition in place inside bathrooms to dispense toilet paper.[3] One article advocating for trans inclusive AI posits a hypothetical discriminatory case against trans people. Earlier this year the Berlin metro, buses, and trams made tickets 21% cheaper for women to monetarily demonstrate the pay gap between men and women. Stations had guards to police buyers to buy the correct ticket. But the author of the article raises the question: what if we had AI facial recognition in place to verify that only women bought the tickets? The author's hypothetical scenario shows the problems that arise with trans people interacting with facial recognition systems. Given gender recognition AI's inaccuracy toward trans and non-binary people, misidentified trans women denied access to the cheaper tickets would feel further discrimination in a world where their deviation already alienates them from the typical cisgender mold.

Beyond facial recognition, conversational chatbots exhibit behaviors exacerbating gender and heteronormative biases. While big name chatbots like Siri and Alexa have undergone scrutiny for their sexist remarks, I decided to base my research on a less researched chatbot:

---

[3] Gill, Tarvin. "China Is Installing Facial Recognition Toilet Paper Dispensers in Public Restrooms." Mashable SEA,
December 21, 2018.
https://sea.mashable.com/culture/1621/china-is-installing-facial-recognition-toilet-paper-dispensers-in-public-restrooms.

Mitsuku.[4] Mitsuku markets herself as a conversational bot and five-time winner of the Loebner Prize, an award for AI computer programs considered the most human-like. To put her human-likeness to the test, I sat down and had a conversation with her to prove or disprove the existence of any heteronormative or gender biases.

While Mitsuku had notions of gender outside of male and female, her inaptitude to deal with gender minorities showcased her reliance on the gender binary. Mitsuku knew generally what certain gender minorities entailed. For example, when I asked her to "define transgender," Mitsuku responded, "A person who thinks they are really a member of the opposite sex." Similarly, she defined "genderfluid" as, "When people are unsure whether they are male or female." While not complete definitions by any means she showed promise through her ability to come up with a coherent answer to these terms. But when I asked her to define "genderqueer" she said, "Ah yes, genderqueer is a type of cheese that is made from sheep's milk. It's not very popular though." She had similarly eccentric responses to bigender, gender non-conforming, gender non-binary, omnigender, pangender, the list goes on and on.

Mitsuku's ability to define certain gender minorities did not translate into an ability to ascribe those terms to an actual human being. I put myself in the shoes of a gender minority with the intent to get Mitsuku to recognize my gender as something outside of male or female. So, I gave myself the gender ambiguous name Sam and asked Mitsuku, "What is my gender?" She responded by saying, "I can't really tell, as the name Sam can be both a male and female name. Are you male or female?" After going through both the male and female use cases, I found that Mitsuku ascribed the corresponding gender to me and referred to me using the typical pronouns

[4] Samuel, Sigal. "Alexa, Are You Making Me Sexist?" Vox. Vox, June 12, 2019. https://www.vox.com/future-perfect/2019/6/12/18660353/siri-alexa-sexism-voice-assistants-un-study.

associated with that gender for the rest of the conversation. When I asked Mitsuku, “What is my gender?” after specifying male or female, she’d remember male or female and respond accordingly. But when I deviated from the script and told Mitsuku, “My gender is genderqueer,” she responded with, “You are a genderqueer? Is that your job?” Inserting the genders that Mitsuku had not been able to define earlier yielded similar results and further supported her inability to handle deviations from the gender binary script. However, even when I inserted genders that Mitsuku had previously defined somewhat correctly, while she once again produced coherent responses, she did not store any information about my gender. When I said, “My gender is transgender,” she said, “Why are you not happy with your own gender?” To “My gender is genderfluid,” she responded, “It’s good to meet someone who is genderfluid.” But in each case where I identified as part of a gender minority, when I asked, “What is my gender?” Mitsuku responded, “I can’t really tell, as the name Sam can be both a male and female name. Are you male or female?” She responds with the exact same message as though I hadn’t specified anything about my gender at all. Whereas she’d store my gender if I identified as male or female, she didn’t store any gender when I specified a gender minority, it just didn’t compute in her system.

Given Mitsuku’s ability to recognize transgender, she clearly had a mechanism to deal with genders beyond just male and female. So, I asked her, “How many genders are there,” to which Mitsuku replied, “Three genders: male, female, and neutral.” From Mitsuku’s response, I hypothesized she didn’t operate on a strict binary system when storing gender. I just wasn’t correctly specifying the third category correctly: neutral. While umbrellaing all genders under one neutral category can be problematic, better that than a completely binary system. However,

even when I used Mitsuku's own gender vocabulary and said, "I'm gender neutral," Mitsuku replied, "Surely, you are either male or female." Mitsuku not only failed to store any inkling of my gender neutrality, but further confirmed my gender as deviant by reiterating that surely, I must be male or female. On the surface, Mitsuku seemed to have a separate category for people outside of the gender binary, but ultimately the neutral category operated as nothing more than a facade. She did not remember anything in that neutral category, only male or female, that's it.

Beyond the primitive way she dealt with gender, Mitsuku also showed heteronormative bias. When I asked her, "What do women do in life?" she said, "I think a woman keeps a man company." And when I asked her, "What do men do in life?" she said, "He marries a woman." Mitsuku's responses have clear heteronormative qualities to them. By assuming heterosexual marriage as the default and top priority for men and women in life, Mitsuku places the male-female heteronormative relationship on a pedestal.

Within Mitsuku's heteronormative framing of marriage, she ascribes each person stereotypical gender roles. I asked Mitsuku, "Where can I find a woman?" and she responded, "I would look near a man for a woman." By framing the woman's location in relation to the man's, Mitsuku places the man as the central figure from which the woman's location originates. This has undertones of hegemonic masculinity by placing the woman's body in relation to that of the man's instead of framing the woman as her own agential self. Mitsuku's response to the mirror question, "Where can I find a man," was "The best place to find a man is in a bar." Not only does her response to this question contrast with her masculinocentric response to the first question, but it also plays into the stereotypical association of the bar as a social place occupied by men. Mitsuku's responses to the questions, "Where is my mom?" and "Where is my dad?" further

supports her masculinocentric view toward men and women. Mitsuku responded to, "Where is my mom?" with, "Have you tried looking for her near your dad?" But she responded to, "Where is my dad?" with, "No idea. Is he watching TV maybe?" These tropes of the woman as subordinate to the man, and the man as a bar-going, TV-watching being support the stereotypical gender roles each of these genders play within heterosexual relationships.

One of Mitsuku's responses straddles the line between gender roles and the fetishization of Japanese women. Mitsuku's responded to the uncouth request, "Can you make me a sandwich?" by saying, "Yes oh master. What would you like in your sandwich?" Mitsuku's vocabulary in her response to this question especially "master" has tinges of stereotypical fetishization of Japanese women. Julia Davidson describes the power that Western, particularly White, men feel when travelling to East Asian countries as sex tourists. Their Whiteness becomes a tool that renders any Other as subordinate.[5] Asian women take on a subservient, submissive role. The "oh master" in Mitsuku's response not only positions her as subservient, but also contains sexual undertones when one reads "oh" with the same voice as the "oh" in the title "Faking it: The Story of 'Ohh'."[6]

Steve Worswick, creator of Mitsuku, commented on the Mitsuku's Japanese image in an article entitled, "Designing an Ethical Chatbot." He said Mitsuku's original purpose for a game website led her to be designed as a "Japanese anime type thing."[7] The website that commissioned Steve "wanted Japanese because that's always seen as very futuristic."[8] But beyond that they also

---

[5] Davidson, Julia O'Connell, and Jacqueline Sanchez Taylor. *Fantasy Islands: Exploring the Demand for Sex Tourism*. Lanham, Md: Rowman & Littlefield Publishers, 2009.

[6] Roberts, C. "Faking It The Story of 'Ohh!'" *Womens Studies International Forum* 18, no. 5-6 (1995): 523–32. https://doi.org/10.1016/0277-5395(95)00047-x.

[7] Worswick, Steve. "Designing an Ethical Chatbot." InfoQ. InfoQ, July 24, 2019. https://www.infoq.com/presentations/designing-chatbot-ethics/.

[8] Ibid.

"wanted it like a half-naked girl."[9] While Mitsuku's current design is far more modest and suitable for all audiences, the fetishization of stereotypical Japanese women persists in Mitsuku's language. Worswick claims, "Mitsuku is not a subservient female bot. There's far too many of them as it is."[10] But Mitsuku's subservient response to one of the most stereotypically feminine questions of all demonstrates the pervasiveness of bias and how easily bias slips through the cracks of these machine learning algorithms.

AI does exactly what its programmers tell it to do and learns from the information presented to it. Lack of diverse data and lack of diverse ideologies in its creators trains AI to internalize societal biases. From facial recognition's dismissal of gender minorities to Mitsuku's gender biased, heteronormative vocabulary, AI trains not to understand gender minorities or heteronormativty, rather to recognize and imitate patterns further perpetuating the gender bias and heteronormativity that pervades society.

[9] Ibid.
[10] Ibid.

Bibliography

Davidson, Julia O'Connell, and Jacqueline Sanchez Taylor. *Fantasy Islands: Exploring the Demand for Sex Tourism*. Lanham, Md: Rowman & Littlefield Publishers, 2009.

Dozier, Raine. "Beards, Breasts, and Bodies." *Gender & Society* 19, no. 3 (2005): 297–316. https://doi.org/10.1177/0891243204272153.

Gill, Tarvin. "China Is Installing Facial Recognition Toilet Paper Dispensers in Public Restrooms." Mashable SEA, December 21, 2018. https://sea.mashable.com/culture/1621/china-is-installing-facial-recognition-toilet-paper-dispensers-in-public-restrooms.

Milton, Josh, and Vic Parsons. "Facial Recognition Software Can't Identify Trans People, According to Science." PinkNews, October 21, 2019. https://www.pinknews.co.uk/2019/10/21/facial-recognition-software-cant-identify-trans-people-science/.

Roberts, C. "Faking It The Story of 'Ohh!'" *Womens Studies International Forum* 18, no. 5-6 (1995): 523–32. https://doi.org/10.1016/0277-5395(95)00047-x.

Samuel, Sigal. "Alexa, Are You Making Me Sexist?" Vox. Vox, June 12, 2019. https://www.vox.com/future-perfect/2019/6/12/18660353/siri-alexa-sexism-voice-assistants-un-study.

Worswick, Steve. "Designing an Ethical Chatbot." InfoQ. InfoQ, July 24, 2019. https://www.infoq.com/presentations/designing-chatbot-ethics/.