"Dear Google; Love, Grandma":
Virtual Personal Assistants and the Political Implications of Humanizing Algorithms

Xinlan Emily Hu

20 May 2019

"Dear Google," the elderly woman types into the search bar, pushing up her spectacles to the brim of her nose. "Could you please tell me where I can buy a gift for my grandson?"

To a technology-fluent young person, such an interaction with Google is inappropriate and comical. But this brief anecdote illustrates more than just a generation gap. It also provides insight into a broader point: in a world inundated with artificially intelligence agents, it is all too easy to confuse the human with the non-human.

This essay explores our interaction with algorithms and artificial intelligence, drawing from virtual personal assistants as a primary (but not exclusive) case study. In the first half, I trace the origins of trust in AI agents, arguing that virtual personal assistants make it difficult for users to question their authority or apply scrutiny to their services. Specifically, virtual personal assistants use two important tactics: first, they convincingly present themselves as human-like beings who seek our friendship; and second, they appear to be objective, truthful tools, even when this is not always the case. In combination, these two features generate an unprecedented amount of trust in machines.

In the latter half of this essay, I outline a framework for evaluating what corporations do with our trust. Rather than broadly criticize all corporate data collection, I argue for a distinction between legitimate and illegitimate information processing using a standard of overreach. I develop my framework with an analysis of two threats that AI poses to democracy. Ultimately, complacency in trusting AI agents not only puts users personally at risk, but also enables powerful entities like corporations and governments to use technology for unethical purposes. Users, engineers, and policymakers alike must not only be aware of their own personal ethical responsibility, but also of their role in a delicate information system.

The first notable aspect of a virtual personal assistant is its perceived humanness. In social science, the Computers as Social Actors (CASA) paradigm[1] posits that users interact with technology in the same manner as with humans, despite recognizing the objects' non-humanness. Although CASA applies very broadly, the effect is particularly salient in our case study of virtual personal assistants. AI assistants heighten the CASA effect by design: devices such as the Amazon Echo or the Google Home are intentionally socially interactive. The agents are consciously named, gendered, and anthropomorphized; they are playful and intelligent. These features are design affordances that nudge users to give a non-human entity human treatment[2].

Moreover, having a personable AI agent is simply good for business. A study from Cornell University analyzed 851 Amazon.com reviews of the Echo device and found that satisfaction with "Alexa" (the agent's named persona) strongly predicts satisfaction with the

[1] Purington, Amanda, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. "'Alexa Is My New BFF'." (*Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA 17*, 2017. doi:10.1145/3027063.3053246), 2.

[2] Ibid.

device itself, regardless of technological issues with the hardware[3]. In other words, corporate incentives skew in favor of a more human-like agent—a "person" whom users can trust and develop a loyalty that blinds them to the product's other flaws.

Despite presenting themselves as friendly and human, AI agents also do not shy away from an air of scientific authority. The Google Home, for instance, is simply invoked with "Hey Google"—with no additional human name—perhaps because the word "Google" has so strongly associated itself in modern vernacular with the searching algorithm. Microsoft's Cortana is named after a robot character in the video game *Halo*; Amazon's Alexa is named after the ancient Alexandria library[4]. Virtual personal assistants appear human, but also leave their machine-ness exposed.

This second essential feature of virtual personal assistants (that they build trust through a sense of objectivity) is achieved through such homages to robotics, science, and knowledge. To acknowledge that these agents are made of algorithms lends them a unique legitimacy: they are "mathematical, logical, impartial, consistent. … Conclusions described as having been generated by an algorithm wear a powerful legitimacy, much the way statistical data bolster scientific claims[5]." The fact that these agents are algorithm-powered thus lends them a second form of authority: not only are they trustworthy because they are our friendly advocates, but their decisions must also be unquestionably correct, for an algorithm's judgement is far superior to that of a mere human mind.

Combined, these two features of virtual personal assistants generate a powerful amount of trust—and with it, the user data of millions. They provide wake-up calls; they order gifts; they recommend Italian restaurants for date night. But trust can blind users to the risk of allowing algorithms to micro-manage their lives. Even the most well-intentioned of technologies, if given unchecked to the powerful, becomes an undemocratic force. In particular, virtual personal assistants bring into light two types of threats: threats to privacy and security, and threats to information access.

Privacy and security, the first threat, have become a recent hot topic. My approach to this topic takes a direction distinct from other scholars. Rather than criticize all corporate data collection categorically, I argue for drawing a line between legitimate and illegitimate uses of data. My initial, loose definition—to be slowly refined in the course of this section—is that illegitimate uses of data *overstep the bounds* of the corporation's reasonable limits of influence. To give a rough intuitive sketch of this claim, consider Spotify. It seems reasonable to allow Spotify to collect data about your favorite music, the better to recommend songs that you might enjoy in the future. It seems unreasonable for Spotify to pass along your listening history of political podcasts to a potential employer, the better to discriminate against your personal beliefs.

Contrast my claim with that of John Cheney-Lippold, who criticizes all corporate use of data. Cheney-Lippold claims that these invasions of privacy are means of "collecting enough

[3] Ibid., 6.

[4] LaFrance, Adrienne. "Why Do So Many Digital Assistants Have Feminine Names?" (The Atlantic. March 30, 2016. Accessed May 23, 2019. https://www.theatlantic.com/technology/archive/2016/03/why-do-so-many-digital-assistants-have-feminine-names/475884/).

[5] Gillespie, Tarleton. "Algorithm." in *Digital Keywords: A Vocabulary of Information Society and Culture*, ed. Benjamin Peters. (Princeton, NJ: Princeton University Press, 2016), 23.

data to grow from mere search engines into much more profitable advertising behemoths, capable of providing query results next to exactly defined commercial propaganda[6]." He takes a Marxist approach in criticizing data collection: in turning users into means of production, he notes that corporations have flattened the world into abstract, hyper-specific, marketable categories: "high-income," "married with children who live in fashionable homes on small, manicured lots," and holding an "advanced degree with sophisticated tastes to match their credentials[7]."

My approach is not so critical of data collection and advertising on principle. I focus instead on practical threats to publics. In other words, there is nothing inherently wrong with using data to generate interesting content, nor is there anything inherently distasteful—aside from the name—about a "profitable advertising behemoth." After all, advertisement practices have existed since the 1950's[8], and given that capitalism is unlikely to change, it is perhaps slightly more pleasant to see a relevant ad than a generic one.

The trouble occurs when information generated from data collection seeps beyond the original platform. Digital trails can reveal "potentially sensitive information such as religious or political beliefs, sexual orientation, race or ethnicity, physical or mental health status, or sex or gender identity;" one's Facebook group memberships can, with a strong degree of accuracy, predict whether he or she is gay[9]. And all of this data can be easily linked to both public and private datasets, such as an individual's voter registration[10]. From there, those who have been "outed" as gay or suspected to have health problems can suffer political disenfranchisement, employment discrimination, or inflated health care premiums. Worse, this discrimination can exist outside of legal protections because it is based merely on statistical inference. Facebook successfully defended its ability to infer sensitive details such as race and sexual orientation by claiming that "they have no intention to infer sensitive traits, but only to assume an affinity[11]." A reasonable act of data collection (providing interesting content) thus suddenly becomes an unacceptable overreach.

Initially, the "overstep" seems innocent: a company makes a bit more revenue by selling data to another party. But once sold to the third party, the data's integrity is lost—personal information can lie exposed for unprivileged eyes to see.

Gina Neff and Dawn Nafus articulate the concept of *contextual integrity,* arguing that even when identifying information is stripped away, "what makes data 'private' is the lineup of people or institutions 'in context' for issues related to the data and the body to which the data refers[12]." Thus, selling data is an overstep by calculation. The sale breaches the users' expectation that their information will remain within the confines of the product, for "it is one thing for companies to have access to personal information in order to be able to deliver a

[6] Cheney-Lippold, John. *We Are Data: Algorithms and the Making of Our Digital Selves*. (New York, NY: New York University Press, 2019), 22.

[7]Ibid., 56.

[8] Wachter, Sandra. "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising." (May 15, 2019. Accessed May 22, 2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3388639), 4.

[9] Ibid., 8.

[10] Ibid., 10.

[11] Ibid., 12.

[12] Neff, Gina, and Dawn Nafus. *Self-tracking*. (Cambridge, MA: MIT Press, 2016), 35.

useful product, but it is quite another for that company to then also sell that data. …[T]hat data is still a part of ourselves[13].” Neff and Nafus use strong language; they frame the loss of contextuality as a physical violation of the user.

Contextual integrity is a useful lens for our framework: it explains the salience of the aforementioned examples. To have your information contained in-app feels natural; to have it released to the world feels as though a friend has betrayed your trust.

There are even cases when virtual personal assistants are actively negligent with users' trust. For example, each device stores a significant amount of personal data (payment information, home addresses, birthdates, smart appliances)[14], much of which is not in fact stored properly. In the case of Amazon Alexa, for instance, all of these interactions takes place through a client-server structure known as an Alexa Skill. But in a 2017 analysis of 11,827 Skills, 75% of Skills lacked a privacy policy[15]. And the device itself is shockingly vulnerable to brute-force attacks[16], phishing[17], and ultrasonic commands that can wake the device with sounds inaudible to humans[18]. All of these security risks point to a need for greater scrutiny on the part of both consumers and lawmakers.

The second threat to democracy—controlling information access— is slightly more complicated to define within the “overreach” framework. That is, information-based corporations generate revenue from serving as successful gatekeepers: Google, for example, points users to YouTube (which it owns), rather than to Vimeo or Youku[19]. This aspect of the product has the very purpose of directing users to outside resources and experiences, so it does not make sense to say that influencing users' actions outside of the bounds of its service constitutes an “overreach.”

In this case, the following litmus test determines the bounds of reach: *did the company's service work as expected?* In other words, the expectations define the boundaries. If, say, the user reasonably expects accurate, objective search results, the results should—to the best of the company's ability—be truly objective. Otherwise, just as advertisements on Instagram require a tag of “#ad” or “Sponsored,” users should be informed that they may not be consuming the content they expect.

At present, corporations instead present non-objective data as pure fact. Scholar Safiya Noble points out, among many examples, that queries for “Black girls” or “Black women” result in highly sexualized content, and that, during Obama's time in office, “Ni**er house” redirected

[13] Ibid.

[14] Haack, William, Madeleine Severance, Michael Wallace, and Jeremy Wohlwend. “Security Analysis of the Amazon Echo.” (18 May 2017. Accessed 22 May 2019. https://pdfs.semanticscholar.org/35c8/47d63db1dd2c8cf36a3a8c3444cdeee605e4.pdf), 1-3.

[15] Alhadlaq, Abdulaziz, Jun Tang, Marwan Almaymoni, and Aleksandra Korolova. “Privacy in the Amazon Alexa Skills Ecosystem.” (Star 217: 11. https://www.petsymposium.org/2017/papers/hotpets/amazon-alexa-skills-ecosystem-privacy.pdf), 1.

[16] Haack, et al., “Security Analysis,” 6.

[17] Zhang, Nan, Xianghang Mi, Xuan Feng, XiaoFeng Wang, Yuan Tian, and Feng Qian. “Understanding and Mitigating the Security Risks of Voice-Controlled Third-Party Skills on Amazon Alexa and Google Home.” (29 June 2019. Accessed 22 May 2019. https://arxiv.org/abs/1805.01525) 2.

[18] Zhang et al., “Understanding and Mitigating,” 1.

[19] Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. (Harvard University Press, 2015), 9.

to a search for the White House[20]. Researchers have also found that Google Search displays higher-paying jobs for men than for women, and have served young LGBT-identifying individuals gay conversion advertisements[21]. If users interpret these results as if they were objective, they would learn a deeply skewed version of reality. As Noble writes, "information assumed to be 'fact' (by virtue of its legitimation at the top of the information pile) exists because racism and sexism are profitable[22]." Algorithms are fed by profit-motivated corporations, with data that is skewed and non-racially representative[23]. So goes the tech industry aphorism: "garbage in, garbage out."

For those who seek to be informed citizens, for those who seek to participate in the public sphere, for those who seek social justice—algorithmic gatekeeping can be stifling. The case study of virtual personal assistants represent a unique extreme: because of the client-server nature of their architecture[24], each query returns exactly one response. Users have no opportunity for further questioning. They can only rely on the trust and authority of their personal assistant—and the black box of proprietary algorithms powering its "mind."

Algorithms may see their content as mere bits and bytes. But algorithms exist within the confines of society, causing them to become unintentionally political. Unlike normal discourse, the political side of algorithms do not result from rational debates in the public square or from a passionate place of personal conviction. They are parameters of training data, representing minute distinctions[25] that may not even be fully intelligible to their creators. But in a way, they do arise from the same publics; they are a product of social relations. With each passing day, "the Internet is both reproducing social relations and creating new forms of relations based on our engagement with it. …As users engage with technologies such as search engines, they dynamically co-construct content and the technology itself[26]."

So we must not see algorithms as superior and more objective than humans. Rather, we must see them as reflections of an imperfect society. As users and citizens, we deserve better. We deserve an understanding of what to expect from the technology we use—whether it is being used to make inferences about our lives; whether it keeps our data secure; whether its delivered product in fact meets our expectations. Regulation should ensure that technology companies do not overstep their bounds, and consumers, for their part, should be skeptical before giving out their trust. For example, Sandra Watchter's suggestion of a "right to reasonable inferences" provides legal protection to users when their data is used to infer sensitive information[27]. Further controls about how and why data can be sold will be another step in the right direction.

And in the end, should we tell Google "please" and "thank you?"

Well, hey Google, check your sources. Please?

[20] Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. (New York: New York University Press, 2018), 4-7.

[21] Wachter, "Affinity Profiling," 9.

[22] Noble, *Algorithms of Oppression*, 31.

[23] Gillespie, "Algorithm," 21.

[24] Haack, et. al, "Security Analysis," 2.

[25] Gillespie, "Algorithm," 21.

[26] Noble, *Algorithms of Oppression*, 151.

[27] Wachter, "Affinity Profiling," 30-31.

Bibliography

Alhadlaq, Abdulaziz, Jun Tang, Marwan Almaymoni, and Aleksandra Korolova. “Privacy in the Amazon Alexa Skills Ecosystem.” Star 217: 11. https://www.petsymposium.org/2017/papers/hotpets/amazon-alexa-skills-ecosystem-privacy.pdf.

Cheney-Lippold, John. *We Are Data: Algorithms and the Making of Our Digital Selves*. New York, NY: New York University Press, 2019.

Gillespie, Tarleton. “Algorithm.” in *Digital Keywords: A Vocabulary of Information Society and Culture*, ed. Benjamin Peters. Princeton, NJ: Princeton University Press, 2016.

Haack, William, Madeleine Severance, Michael Wallace, and Jeremy Wohlwend. “Security Analysis of the Amazon Echo.” 18 May 2017. Accessed 22 May 2019. https://pdfs.semanticscholar.org/35c8/47d63db1dd2c8cf36a3a8c3444cdeee605e4.pdf.

LaFrance, Adrienne. "Why Do So Many Digital Assistants Have Feminine Names?" The Atlantic. March 30, 2016. Accessed May 23, 2019. https://www.theatlantic.com/technology/archive/2016/03/why-do-so-many-digital-assistants-have-feminine-names/475884/.

Neff, Gina, and Dawn Nafus. *Self-tracking*. Cambridge, MA: MIT Press, 2016.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.

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

Purington, Amanda, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. ""Alexa Is My New BFF"." *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA 17*, 2017. doi:10.1145/3027063.3053246.

Wachter, Sandra. “Affinity Profiling and Discrimination by Association in Online Behavioural Advertising.” May 15, 2019. Accessed May 22, 2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3388639.

Zhang, Nan, Xianghang Mi, Xuan Feng, XiaoFeng Wang, Yuan Tian, and Feng Qian. “Understanding and Mitigating the Security Risks of Voice-Controlled Third-Party Skills on Amazon Alexa and Google Home.” 29 June 2019. Accessed 22 May 2019. https://arxiv.org/abs/1805.01525.