"Alexa, Steal Our Data!" (click this link!)

Background

Amazon's Alexa is a voice assistant feature integrated into many of their products, and has even become involved with third-party products as well, enabling the creation of a smart home. Alexa has the ability to complete a variety of simple and complex tasks, such as dimming lights in a house, locking doors, adjusting thermostats, playing music, or describing current weather (Rawes). Other companies have created voice assistants similar to Alexa, like Apple's popular Siri and Google Home's Google Assistant. To narrow the scope of this assignment, our group decided to focus on Amazon, assessing the benefits and costs of Alexa.

Though Alexa's abilities may seem magical, users are actually interacting with a cloud-based service, called the Alexa Voice Service (AVS). AVS is designed to convert voice recordings into commands that Alexa can interpret and complete. Alexa is unique because it is a fully programmable service that can complete a variety of different tasks (Baguley). To interact with Alexa, users must state her name before the task they want completed. Amazon claims that Alexa isn't listening before her name is used, however, research has found otherwise. While Alexa offers a variety of benefits, her service comes at a cost: personal data.

How Amazon's Data Collection Provides Value to Amazon & Consumers

Soft Sell Marketing (Benefits the Consumer, Amazon, and Amazon Partners)

Amazon is a leader in collecting, storing, processing, and analyzing personal information from its customers as a means of determining how customers are spending their money. Alexa has the ability to distinguish different users based on voice tone and pitch, allowing the device to create a personalized experience depending on the user (Schwartz). When users visit Amazon,

they receive recommendations based on previous history and personal data for future products to buy. More importantly, however, Alexa is also a strong channel through which companies are starting to advertise their products. Currently, Alexa has a feature called the Zyrtec "skill" which allows users to receive information on allergens in the air when asked. However, this feature advertises for the brand Zyrtec inadvertently, a way for Amazon to advertise without the users necessarily feeling like they are being manipulated (Amazon). This soft-sell marketing attempts to ensure that users don't feel targeted by the data used to determine their interests.

Proactive Marketing (Benefits the Consumer, Amazon, and Amazon Partners)

According to Rohit Prasad, Amazon's head scientist, Amazon plans to move from passive to proactive interactions to better assist their users, while integrating advertising for some products in that process (Hao). For example, Amazon hopes to be able to help users plan a night out. If a user asks Alexa to buy them a movie ticket, she will then follow up and ask the user if they would also like her to book an Uber (advertising paid for by Uber). Once enough users have booked an Uber following a night out, Amazon will use that frequency of data to change Alexa's reaction to booking a night out. She will recommend ordering an Uber from the start, which shows how Amazon hopes to anticipate users requests, instead of waiting for them to be given. This level of prediction is impossible without the data Alexa collects from her users, and because of the collected data, Amazon knows they are advertising something their users are likely not to feel targeted by, based on the amount of users who do in fact book an Uber after a movie (Hao)

The result of using voice assistants and AI in advertising will lead customers to rely on the suggestions of their voice assistant, rather than finding their own trusted brands. It will change customer acquisition, customer satisfaction, and retention (Dawar). In the case of Alexa,

Amazon uses what is called predictive analysis, using existing data sets from user behavior on their platforms to extract information to identify trends in behavior, target their marketing strategies, increase customer satisfaction, and build customer loyalty by advertising the perfect product to the correct customer (Wills). Finally, Amazon would also be able to retain their customers through their voice assistant, because their collected data would be able to identify better suggestions of products to give that customer. Amazon can also advertise their own brand, and this is key to their advertising strategy. While some of these incredible features of Alexa are yet to come, and benefit both Alexa and the customer, they are reliant on the data Alexa collects.

Impacts of Amazon's Data Collection

AI assistants like Alexa have many economic benefits, especially in a world with an increasing number of choices. However, Amazon confirms they do not always delete Alexa data even if a user deletes audio files, and they supposedly store that data for machine-learning and feature improvement purposes without notifying the user. Even so, this leads to the debate of whether sacrificing privacy is ethical even if it appears to benefit the consumer (Kelly and Statt). Lawsuits have been filed against Amazon as consumers realize that voiceprints containing private details of millions of users, including children, are stored, and in some cases used as criminal evidence (Bronstad). If Alexa is listening and recording all the time, is consumer privacy obsolete? Is this lack of privacy something most consumers are willing to sacrifice for convenience? Nadler, Crain, and Donovan's "Weaponizing the Digital Influence Machine" discusses these privacy issues in the context of political manipulation, stating that digital influence machines employ weaponization to target weak points where groups and individuals are most vulnerable to strategic influence (Nadler et al. 6). In the same way, Amazon's Alexa

targets customers on psychological and emotional terms to gather consumer profiles and manipulate purchasing behavior.

Perceived Benefits of Amazon Alexa's Data Collection

Benefits for the Elderly and Disabled

Alexa's role as a voice assistant has the ability to assist various groups of people such as the elderly. Specifically, Alexa can help those suffering from Parkinson's disease to continue to live on their own. For example, Allie's grandma has difficulty remembering to take her medication and go to bed at a normal time, and as her disease progresses, it becomes more important for her to follow a routine. Due to the side-effects of shakiness, she also has difficulty completing many tasks such as turning off the TV or turning off the lights. By placing three Amazon Echos throughout Allie's grandma's house, they are able to set alarms through Alexa to keep her grandma on track, thereby ensuring she will take her medication at the correct time, and control the lights and TV through voice commands.

Alexa also provides an extra layer of support in case of an emergency. Alexa can easily call 911, creating a sense of comfort for Allie's family. Overall, Alexa has given Allie's grandma the tools to manage her Parkinsons and give her a sense of relief that she always has a way to call for help. Allie's grandma isn't the only one benefitting from Alexa, as those suffering from dementia or other aging diseases have received aid through the voice assistant (Cemental).

Diagnosing PTSD in War Veterans

Published online April 22, 2019 in the journal Depression and Anxiety, a study found that an artificial intelligence tool has 89% accuracy of distinguishing people with and without PTSD using their voice. While the study is not about exploring the disease mechanisms behind PTSD,

the theory is that traumatic events change brain circuits which process emotion and muscle tone, which affects a person's voice. "Speech is an attractive candidate for use in an automated diagnostic system, perhaps as part of a future PTSD smartphone app, because it can be measured cheaply, remotely, and non-intrusively," says lead author Adam Brown (NYU Langone Health / NYU School of Medicine). Plans for the future are to continue improving and validating the tool using more data.

Data collected from talking to Alexa offer enough information that can be used for diagnosis. If companies were allowed to analyze this data they would be able to relate the user experience of Alexa to support dealing with PTSD. This could ease the struggle of individual dealing with PTSD and also improve the artificial intelligence tool accuracy in distinguishing voices with PTSD.

Improving Mental Health

By tracking command history and listening to the mood and physical activity of the user, Alexa can provide support for the user's mental health. In early 2019, Amazon announced that Alexa became HIPAA compliant, meaning third parties can create voice tools to leverage and transfer medical data with extra security protection (Ross, Casey, et al.). Jen Heape, one of the creators of Mindscape, an in-home therapist tool, stated, "We don't capture any data on individual users at all. But we can see what the trigger words are, we can see the topics that people say that they are stressed about most" (Lock). Thus, the application can add new discussions about the topics that trouble a large number of users without compromising their privacy. Active engagement with the tool may improve the mental health of anyone, and improvements to the artificial intelligence will result in interactions feeling more natural.

Perceived Downsides of Amazon Alexa's Data Collection

Lacking Privacy

In Amazon's privacy policy, it states that Alexa collects personally identifiable information, though they do not attempt to analyze the information to determine the identity of the user. Alexa gathers personal information including user data, searches performed, website visits, advertisements viewed, IP address, frequency of website visits, website cookies, and other information that may be collected through a third party skill. Alexa shares the data with Amazon and the subsidiaries it controls as well as any personally identifiable information with third parties. According to the privacy policy, "the information we learn from users helps us personalize and continually improve your Alexa experience and provide information about Internet trends, website popularity and traffic, and related content" (Alexa Internet - Privacy Notice). Having personally identifiable information collected and sold can give companies the power to target users individually. Though the company states that they do not attempt to analyze the personally identifiable information, there is no documentation stating that third party companies will not analyze the information.

Corrupt Business Practices

When a company's primary motive is to make money, its privacy and transparency policies never benefit the consumer. As an international company that focuses on cutting costs and time, it is unlikely that Amazon acts on customers' best interests. In fact, Amazon has tried to find other monetization methods past the single cost of purchasing an Alexa-enabled device such as an Echo. To do so, they allowed third party developers to create and sell skills for people to add onto their base Alexa. Amazon will take 30% of the revenue for all third party skills, with

the original developers receiving 70% (Duprey). While this can provide value to users through additional skills, the motivation for this move was finding new revenue streams. Likewise, the addition of third party companies opens the door to potential data security issues. Users may blindly agree to releasing data to the third party that they would have chosen not to if they actually read the privacy policy. The third party may have malicious motivations as well and may try to leverage their access to data to influence people without their knowledge, similar to the events involving Cambridge Analytica and the United States 2016 presidential election. Amazon proved that they care more about their revenue streams than protecting their users from potential crises and abuse of personal data.

Hacking

There have been several instances where users have reported randomly hearing Alexa talking without being prompted. In 2017, Martin Josephson (pseudonym) heard his Amazon Echo dot saying commands based on his previous interactions with the device. It was regurgitating his past commands. There have been many such instances that have turned out to be fairly creepy. Another example is of a man who's echo randomly cheerfully stated "Every time I close my eyes, all I see is people dying" (Lynskey). These odd behaviors can be due to Alexa mishearing words, misinterpreting a command, or being hacked. Many users worry that malicious people can hack into their system and listen in on them, or access past recordings. Though Amazon claims to counter this concern through it's encrypted storage of information, more recently it's been found that it's not so hard to hack these systems (Moynihan). A research team at the University of Michigan studied the system to detect weaknesses in Alexa's security in late 2019. They found that the microphones that activate multiple voice assistants, including

Alexa, can be triggered using lasers, which could allow anyone to hack into someone's device. Because light travels through windows and light can be amplified using various devices, people can access others' Echos from up to 110 meters. This tactic and Echo's connection with other smart devices in a house can empower people to even go as far as triggering a garage door to open so they can steal someone's car. This is a massive loophole in the Echo's design. Amazon was alerted by U-M of this vulnerability so they can attempt to bolster this weakness.

Targeted Advertising

According to Nadler, the Digital Influence Machine is defined as an infrastructure of data collection and targeting capacities that can be used as a form of weaponization when it targets ads to strategically influence vulnerable groups online. Through surveillance and profiling, targeting, and automating and optimizing elements of influence campaigns, a DIM uses 'persuasion' architecture to divide and conquer different social groups (Nadler).

Voice and audio commands show vulnerable aspects in a person. Especially in a product meant to be a companion, targeted advertising based on information collected can be harmful to the user. It's so simple to harvest data to create sophisticated political profiles. Already, Facebook has proven to impact our political opinions after the Cambridge Analytica scandal, imagine the effects of targeted advertising when it comes from an item on which users rely heavily, and treat as their best friend.

Targeted advertisements are a problem politically, but socially they can also have extreme immediate impacts. With extreme amounts of data, companies keep track of values, vulnerabilities, relationships, subconscious thoughts about a user, and much more. For example, Target was able to predict a girl was pregnant, and her father found out through Target's

personalized ads rather than the girl having the ability to tell him (Hill). In this case, there is no recorded evidence of social consequences. However, if this was a surprise pregnancy or an unwanted one, this could have turned wrong. There are other such scenarios where targeted advertisements can be dangerous, like advertisements catered towards someone's identity such as their sexual orientation or gender change before they are ready to tell others. It can be extremely harmful to be outed unexpectedly before the consumer has a chance to prepare for the consequences.

Perceived Outcomes and Reactions to the Benefits/Downsides of Amazon's Alexa

Our group's perspectives regarding the privacy afforded by AI assistants, like Alexa, vary depending on content type, how the data is collected, ad relevance, and company transparency. Voice data currently gathered by these assistants contains emotional data and allows companies to target vulnerabilities in their customers, however most group members felt that surveillance would make most of us uncomfortable and very reluctant to use the device since its invasiveness would yield even more personal information. Arguably, perspectives would not change based on which company is collecting data, since all companies generally practice the same policies. However, we did identify that group members' reluctance to use voice assistance would be eased if they heard that a company's priority for using personal data is to benefit them, instead of simply for company profit. Ad relevance is important too, because targeted ads can help a consumer find new brands, but many group members feel unsettled when they see an ad highly relevant to their recent offline interactions. For this reason, group members would rather see targeted ads based on personal information consensually given to the company.

The majority of users do not read their privacy policies because they are usually inaccessible, time-consuming, and vague. This is true for all of our group members except for one, who is deeply invested in dedicating time to finding out what their personal data is used for. Even if a company created fully accessible and transparent policies, there is no evidence to suggest that they would execute those policies, or that users would still read them if they already use the product. Further, most customers would agree that it matters more how the company treats user data regardless of their transparency, however clearly-outlined policies such as those compliant with the GDPR could help customer awareness surrounding the security of their personal information (Olsen). These policies would need to highlight the purpose, intent, accuracy, and storage of the collected data, and how the company remains accountable and confidential throughout the process in order to improve their relationships with customers who are increasingly wary of the threat of data privacy infractions.

Recommendations

Targeted Advertising

In order for users to feel comfortable with their data being used to generate targeted ads, we suggest that companies like Amazon explain to consumers the ways data is used to target them. For example, if Alexa were to tell users that she suggests a specific brand of organic paper towels because the user frequently buys organic products, the user might feel more comfortable. In addition, allowing for feedback based on the ads Amazon is feeding their users, like a "this ad doesn't apply to me" button, would allow users to feel more in control of the content they are given and feel that their opinion is valued.

Lack of Privacy

To address the concern of privacy, companies selling voice assistants should have an opt in policy to some or all data collection. Because opting out of this process likely hurts the company's product, service, or monetization of data, users would have to bare some of the burden of opting out. For example, if a user allowed their behavioral data to be collected but opted out of location services, that user will enjoy less functionality within the app but enjoy comfort that their private data isn't going into the hands of a corporation.

Corrupt Business Practices

In order to protect consumers from corrupt business practices in the voice assistant industry, both key organizations, like Amazon, and third party developers need to be transparent about how data is transferred between the companies and what their intentions are moving forward with what's been collected. Without this transparency, it is too easy for either companies to make immoral decisions for the greater good of their company, at the expense of the user. State and federal laws may aid in setting guidelines on the ethical implications of personal data used to influence someone's future decisions, or devices that collect but do not properly secure data. Government intervention is crucial so that users receive additional rights and companies will be held responsible if they display a lack of ethical standards.

Hacking

To bolster Echo's security measures, Amazon should create a better detection and notification system to alert the user when their device or some data may have been compromised. This will allow consumers to act on the potential threat as it arises rather than after the breach. Additionally, Amazon should improve their voice recognition system. On one hand, this could patch the huge hole in their security that U-M researchers discovered using lasers to gain access

to the device. On the other, this will also reduce the probability of the Echo misinterpreting a word as the specified wake word, thus comforting users that their device is only listening when it should be.

Other Concerns

Some concerns that were not addressed are behind theory in social control. First, Alexa shows accent bias. Most smart speakers work best with voices from US regions where they are popular (Harwell). This can create a digital divide for users whose first language is not English, or potentially have speech impairments. Second, Alexa, by default, has a female name and female voices that portray obedient, obliging, and potentially flirtatious tones. This reinforces problematic gender stereotypes. An example of this is the apologetic and deflecting responses that the assistants give to insults (Specia). Finally, the presence of Alexa can create a Panopticon Effect. This is the idea that users act differently when they think they are being surveilled, whether or not they actually are (Bentham). With the ever present existence of an object that is listening or surveilling at all times in the intimate homes can cause users to inadvertently act differently. It seems only ethical that users are informed about privacy concerns as well as biases that they may gain from Alexa before they use it.

Alexa, terminate paper in 3…...2…….1

Works Cited

"Alexa Internet - Privacy Notice." *Alexa*, www.alexa.com/help/privacy.

Baguley, Richard. "Appliance Science: Alexa, How Does Alexa Work?" *CNET*, CNET, 4 Aug. 2016, www.cnet.com/news/appliance-science-alexa-how-does-alexa-work-the-science-of-amazons-echo/.

Bronstad, Amanda. "Amazon's Alexa Devices Violate Children's Privacy, Class Actions Say." *The Recorder*, 11 June 2019, www.law.com/therecorder/2019/06/11/amazons-alexa-devices-violate-childrens-privacy-class-actions-say/?slreturn=20191108115607.

Cemental, Michelle. "Using Alexa for Seniors with Dementia." *Caring Senior Service*, Caring Senior Service, 10 Sept. 2019, www.caringseniorservice.com/blog/alexa-helps-dementia.

Dawar, Niraj. "Marketing in the Age of Alexa." *Harvard Business Review*, 24 Jan. 2019 www.hbr.org/2018/05/marketing-in-the-age-of-alexa.

Duprey, Rich. "This Could Be the No. 1 Way Amazon Makes Money With Alexa." *The Motley Fool*, The Motley Fool, 14 Oct. 2018, www.fool.com/investing/2018/10/14/this-could-be-the-no-1-way-amazon-makes-money-with.aspx.

Hao, Karen. "Inside Amazon's Plan for Alexa to Run Your Entire Life." *MIT Technology Review*, MIT Technology Review, 8 Dec. 2019, www.technologyreview.com/s/614676/amazon-alexa-will-run-your-life-data-privacy/.

Harwell, Drew. "It's All Greek: Alexa, Google Home Foxed by Chinese, Spanish Accents." *South China Morning Post*, 24 July 2018, www.scmp.com/magazines/post-magazine/long-reads/article/2156455/amazons-alexa-and-google-home-show-accent-bias.

Hill, Kashmir. "How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did." Forbes, Forbes Magazine, 31 Mar. 2016, www.forbes/.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#509a697a6668 .

Jesse, David. "U-M Researchers Use Laser to Hack Voice-Activated Devices like Amazon Echo." *Detroit Free Press*, Detroit Free Press, 7 Nov. 2019, www.freep.com/story/news/education/2019/11/07/alexa-amazon-echo-hackers-university-michigan/2506211001/.

Kelly, Makena, and Nick Statt. "Amazon Confirms It Holds on to Alexa Data Even If You Delete Audio Files." *The Verge*, The Verge, 3 July 2019, www.theverge.com/2019/7/3/20681423/amazon-alexa-echo-chris-coons-data-transcripts-recording-privacy.

Lock, Helen. "This AI Therapy Skill for Alexa and Google Aims to Help with Everyday Mental Health." *The Ambient*, 10 Oct. 2018, www.the-ambient.com/features/mindscape-alexa-therapy-skill-mental-health-1034.

Lynskey, Dorian. "'Alexa, Are You Invading My Privacy?' – the Dark Side of Our Voice Assistants." *The Guardian*, Guardian News and Media, 9 Oct. 2019, www.theguardian.com/technology/2019/oct/09/alexa-are-you-invading-my-privacy-the-dark-side-of-our-voice-assistants.

McMullan, Thomas. "What Does the Panopticon Mean in the Age of Digital Surveillance?" *The Guardian*, Guardian News and Media, 23 July 2015, www.theguardian.com/technology/2015/jul/23/panopticon-digital-surveillance-jeremy-bentham.

Moynihan, Tim. "Alexa and Google Home Record What You Say. But What Happens to That Data?" *Wired*, Conde Nast, 3 June 2017, www.wired.com/2016/12/alexa-and-google-record-your-voice/.

Nadler, Anthony, Matthew Crain, and Joan Donovan. "Weaponizing the Digital Influence Machine." Data & Society Research Institute (2018).

NYU Langone Health / NYU School of Medicine. "Artificial Intelligence Can Diagnose PTSD by Analyzing Voices." *ScienceDaily*, ScienceDaily, 22 Apr. 2019, www.sciencedaily.com/releases/2019/04/190422082232.htm.

Olsen, Nicole. "The 6 Privacy Principles of the GDPR". *Privacy Policies*, 3 September 2019, www.privacypolicies.com/blog/gdpr-privacy-principles/.

Rawes, Erika. "What Is Alexa, Where Does She Come from, and What Can She Do for You?" *Digital Trends*, Digital Trends, 3 Oct. 2019, www.digitaltrends.com/home/what-is-amazons-alexa-and-what-can-it-do/.

Ross, Casey, et al. "Amazon Alexa Now HIPAA-Compliant, Allows Secure Access to Data." *STAT*, 4 Apr. 2019, www.statnews.com/2019/04/04/amazon-alexa-hipaa-compliant/.

Schwartz, Eric Hal. "Amazon Launches Alexa Skill Personalization Features." *Voicebot.ai*, 1 Nov. 2019, www.voicebot.ai/2019/11/01/amazon-launches-alexa-skill-personalization-features/.

Specia, Megan. "Siri and Alexa Reinforce Gender Bias, U.N. Finds." *The New York Times*, The New York Times, 22 May 2019, www.nytimes.com/2019/05/22/world/siri-alexa-ai-gender-bias.html.

Thomas, Laurel. "Hey, Alexa, Stop Listening to Everything I Say." *University of Michigan School of Information*, 26 Nov. 2018, www.si.umich.edu/hey-alexa-stop-listening-everything-i-say.

Wills, Jennifer. "7 Ways Amazon Uses Big Data to Stalk You." *Investopedia*, Investopedia, 4 Dec. 2019, www.investopedia.com/articles/insights/090716/7-ways-amazon-uses-big-data-stalk-you-amzn.asp.