Analyzing the Actions Companies with Voice Recognition Products Can Take to Protect Users' Personal Information and Privacy

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Nafeisha Tuerhong

Spring, 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

ADVISOR

Hannah Rogers, Department of Engineering and Society

Abstract

As of now, about 500 million people use voice-controlled virtual assistants, with 1.8 billion people expected to do so by 2021, according to a study done by iProspect (Martin, 2017). This situation presents the problem of increased insecurity of customers about their data privacy. In addition, companies are required to prepare for this increase and provide equal or better-quality service. Based on this information, the related parties are identified as the general public and companies with voice recognition products. This STS paper focuses on the actions companies took and other actions they could take to enhance data security. The main STS method used in this research is the Actor Network Theory, along with methodologies of questionnaires and interviews conducted on people who own voice recognition products. Therefore, this report mainly uses questionnaires from online sources and in-person interviews conducted by me as the method to analyze both parties' arguments.

Introduction

The new technologies developed by companies nowadays tend to aim for the convenience of the people. One of the growing fields is voice recognition. Generally, voice recognition is commonly used for voice-controlled virtual assistants on smart devices. As a new but still growing technology, the controversies have also developed with the growth. The main concern concludes to be the security and privacy of personal data. With the customer feedback often relating to security, the companies made several attempts to solve this problem. However, the public was not satisfied, and the problem still remains unsolved. Therefore, this paper uses two types of methods to analyze and attempt to resolve this problem: questionnaires from online sources; in-person interviews conducted on people who own voice recognition products around UVA campus. The online questionnaires introduce the concerns that public have about voice recognition products. On the other hand, the in-person interviews discuss the convenience of having such products. Based on the feedbacks and the Actor Network Theory, the paper will analyze the actions that companies have taken, and the suggested actions they could take to resolve this prolonged controversy.

The Power of Technology

The first voice recognition technology was a tool called Shoebox presented by IBM at the Seattle World's Fair in 1962 (Ramos, 2018). With the size of a shoebox, it could perform mathematical functions and recognize 16 spoken words as well as digits 0-9. Later, in the 1970s, with the substantial support of the United States Department of Defense and its Defense Advanced Research Projects Agency, researchers at Carnegie Mellon University created Harpy, which could recognize 1,011 words. After the organizations developed technologies that could recognize word sequences, companies started to build applications for the technology.

Throughout the 1990s, companies like IBM, Apple, and others developed items that used voice recognition (see Figure 1). Even now, technology companies are working to create increasingly sophisticated technology that will automate more processes and tasks we do throughout the day. Even the virtual assistants are capable of learning new words and tasks (Ramos, 2018).

Figure 1. Timeline of Mainstream Voice Assistants (Ramos, 2018)

Different Types and System Architecture

Voice recognition features can be adopted with different virtual assistant devices. There are intelligent/automated personal assistants. These are software that are capable of assisting people with simple and basic tasks using natural language. They are also capable of going online and searching for an answer to a user's question. There are also smart assistants that refer to

devices that can provide various services by using smart speakers that listen for a wake word to become active and perform certain tasks. In addition, there are virtual digital assistants that are automated software applications or platforms that assist the user by recognizing natural language in either written or spoken form. Finally, there are voice assistants. They are digital assistants that use voice recognition, speech synthesis, and natural language processing to provide a service through a particular application (Ramos, 2018).

The overall system design of voice controlled personal assistants consists of 4 stages: data collection in the form of speech, voice analysis and conversion to text, data storage and processing, and generating speech from the processed text output (Dekate, Kulkarni, & Killedar, 2016). Each of the 4 stages is actant that composes this network where those stages interconnect and affect each other by the Actor Network Theory (Crawford, 2020). According to ANT, these actors all have significant and equal part to play in the system, and they have interests that lead to a creation of a network to produce effects that meet these interests. Therefore, all of the 4 stages are necessary actors to produce the voice assistants. They play an equal part in creating a network to accomplish the final result.

The Problems

More recent updates to voice assistants such as Siri and Google have taught the assistants to predict what users want to know before they are asked, pushing notifications on time about a traffic jam on the way to work. This intention of making our digital lives more convenient, requires sacrificing privacy and security (Waddell, 2016). That sacrifice shows technology determinism which is when the technology drives social progress (Bimber, 1990).

As discussed above, voice recognition requires the device to record, store, and translate the recorded voice into strings of words. This process concerned many as there are sensitive and

confidential words that people would feel uncomfortable leaking out. While waiting for the wake word, smart assistants are always listening (Ramos, 2016). Specifically, Apple users reflected that when a conversation about a topic took place, they would encounter a similar content about that certain topic on their device (Sarwar, 2021). As a response to this encounter, the public thinks that "the machines seldom keep our confidence, as they were designed by people who work for companies that desire or even need our personal information to continue to support their products" (Kelly, 2019). Kelly also stated that:

> The business models for the companies that offer virtual assistant products and services to the public have some elements in common. The input of data involves a low- or no-cost device placed within easy reach of consumers. The output of data is a rich buffet of offerings for third-party data brokers and data processors, and eventually to advertisers and other companies hungry for consumers' personal information -- companies that include insurers, political consultants, and financial services. (Kelly, 2019)

Another concern was that the virtual assistants could not answer the questions asked half of the time, making it hard to trust them with financial information or transactions. From the PwC Consumer Intelligence Series voice assistants questionnaire, one user specified that: "The assistant cannot answer my questions half the time, but I am supposed to trust it to help me with something involving money?" (Hayes & Wagner, 2017). It is evident that the users have trouble trusting this technology in the long run, which gives the companies an opportunity to improve the technology.

For instance, one of the U.K. Government departments' decision on "voice ID" caused similar concern. It was Her Majesty's Revenue and Customs, which is the U.K. government's taxes, payments, and costumes authority of the U.K. government. By the year 2019, HM Revenue & Customs has so far signed up about 6.7 million people to its voice identification service, which acts as an ID to sign into their system (Jones, 2018). This "voice ID" is also known as "voice biometrics", it verifies your identity by analyzing and comparing your voice to

the voiceprint the company has stored in its database. As the companies behind this technology introduced, one voiceprint includes more than 100 unique physical and behavioral characteristics of each individual, such as length of the vocal tract, nasal passage, pitch, accent and so on. Those companies claim that it is as unique to an individual as a fingerprint, and that their systems recognize people if their voiceprint was damaged. Following this affirmation, Lloyds Banking Group also claims that the voice recognition systems are capable of spotting the difference between identical twins. However, in May 2017, their system let BBC reporter Dan Simmons's non-identical twin, Joe, successfully access Dan Simmons's account with his "voice ID". This means that it is possible for someone to have access to other users' private data by having similar voiceprint. Therefore, this information proved the concerns that users have are valid.

According to a report from Microsoft Advertising, "41% of users report concerns about trust, privacy and passive listening" (Olson, 2019). Microsoft surveyed a group of people about their concerns regarding digital assistants, the result is portrayed in Figure 2 below.

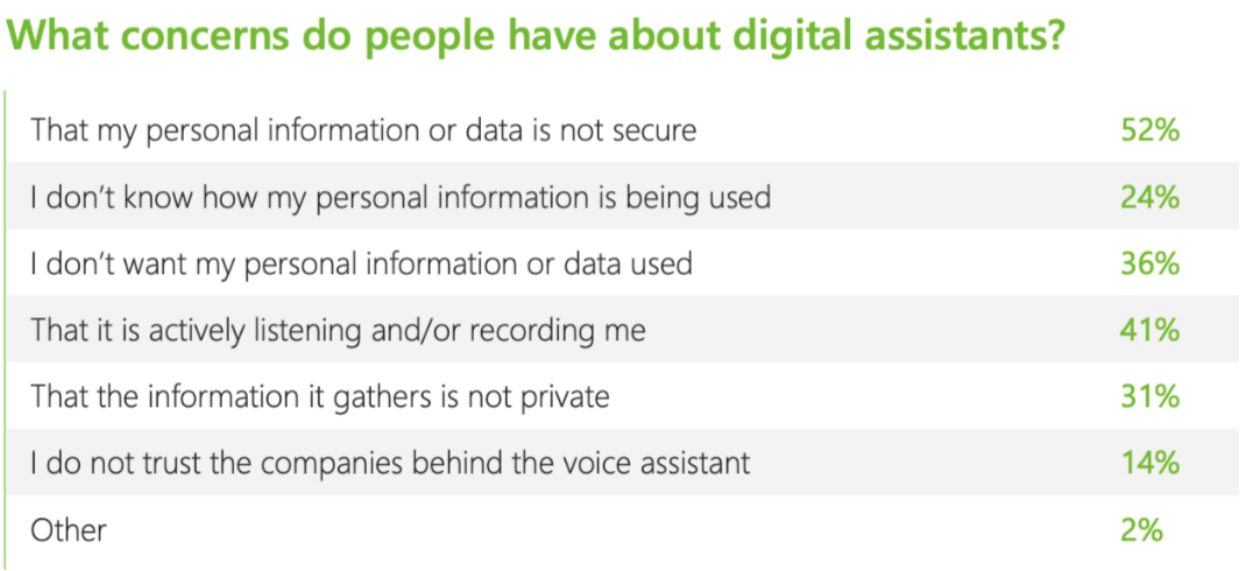

Figure 2. Concerns people have about digital assistants (Olson, 2019)

More than half of the people surveyed think that their personal information or data is not secure with digital assistants. In 2015, one electronics company, Samsung, confirmed this concern with a warning. While keeping the customers updated with the newest features, Samsung warned its customers about discussing personal information near their smart televisions. The warning was targeting the users who use the voice activation feature to control their Samsung Smart TV. It is stated that when this feature is active, their TV listens to its surroundings and has the possibility of sharing that information with Samsung or third parties (BBC, 2015). In addition to the previous problems, it is safe to conclude the public's lack of trust in this technology.

Since the controversies continued for a long time, companies tried to make efforts towards resolving this issue. One of the actions they took was to add an option in the settings for users to select if they want to have personalized advertisements, which would show up based on the users' recent search history with or without the virtual assistant (Google, n.d.). One thing to note is that they did not make any changes to the voice recognition feature. They provided the option to disable personalized advertisements, but the problem of companies storing users' personal information was not resolved.

Privacy Laws

The users of these products are not limited to any age group. Adults, children, and seniors all use different voice recognition products for different reasons. Adults could use voice assistant to ask questions; children could use voice assistant to play games with friends; seniors could use voice assistant to open an application. Those daily activities increase the importance of regulations on data privacy. Comparing to the adults, children and seniors are more likely to

ignore the importance of data privacy and give up sensitive information to others without knowing.

Privacy law in the United States is fairly patchy compared to the world's most comprehensive and powerful data protection law of the European Union (TermsFeed, 2020). One specific case is that we only have California Consumer Privacy Act (CCPA) that has long been a path leader in protecting its residents' privacy and Children's Online Privacy Protection Act (COPPA). It has been suggested that voice assistant technology itself is in fundamental violation of COPPA due to the US's effort in protecting children. On the other hand, the European Union has General Data Protection Regulation (GDPR) that covers almost all commercial activity in the EU that involves the processing of people's personal data. According to ANT, all actants are equally important in the system. They all contribute to the network with their connections (Crawford, 2020). The actants in EU's GDPR network are the groups involved to create such powerful data protection law. The main groups are the companies and public. They have the equal power to create and influence this system. Therefore, they influenced the making of these laws. However, third-party developers and manufacturers must agree with the Terms and Conditions such as collecting any personal information requires a Privacy Policy. For instance, Apple requires all iOS apps hosted in the App Store to be accompanied by a Privacy Policy that must be accessible from within the app itself (TermsFeed, 2020). Apple's reputation for respecting customer privacy depends on third-party developers obeying its rules.

Companies' Actions

For the privacy issues discussed above, "users filed thousands of complaints against Apple, Amazon, and Google for improperly recording and conducting analysis on voice recordings for targeted advertising or software improvement, which sometimes violates specific

states' wiretapping laws" (Cherkassky, n.d.) . When it is pointed out that some of the voice recordings violated the EU's GDPR, the corporations mentioned above made several actions. Based on the ANT, the three major voice recognition product owners interact with the public to create a network that allows both parties to interconnect and affect each other. The relationship between them is necessary to create this network and maintain this system.

Apple

Although Apple did not explicitly unfold this information in its privacy documentation, a small random subset, less than 1% of Siri recordings, are accessed by contractors working for Apple around the world (Hern, 2019b). Part of their job was to grade the recordings and decide if the activation of the Siri was intentional or accidental. Apple claimed that data was analyzed to improve Siri and dictation. However, a contractor working for the company felt uncomfortable with the lack of disclosure to users. The main reason was that the accidental activations of Siri allowed this staff to have access to extremely sensitive personal information about users. Apparently, Siri often mistakenly recognizes the sound of a zip as its wake word. This allowed contractors to listen to recordings of "private conversations between doctors and patients, business deals, seemingly criminal dealings and so on", which were "accompanied by user data showing location, contact details, and app data" (Hern, 2019b).

At first, Apple apologized and suspended the Siri voice grading program that lets contractors have access to the voice recordings from Siri (Cherkassky, n.d.). However, they planned to resume the program after making the following changes:

> First, by default, we will no longer retain audio recordings of Siri interactions. We will continue to use computer-generated transcripts to help Siri improve.
> Second, users will be able to opt in to help Siri improve by learning from the audio samples of their requests. We hope that many people will choose to help Siri get better, knowing that Apple respects their data and has strong privacy controls in place. Those who choose to participate will be able to opt out at any time.

> Third, when customers opt in, only Apple employees will be allowed to listen to audio samples of the Siri interactions. Our team will work to delete any recording which is determined to be an inadvertent trigger of Siri. (Apple Inc., 2019)

Apple provided the opt-in option for users, so they can decide whether they would like their voice recordings to be stored. In addition, they provided the above conditions along with the statement. Based on the conditions, most of the users were satisfied with the opt-out option. Although, it did not provide middle ground for users who want to help Siri get better but afraid of leaking sensitive information.

Amazon

Similarly, Amazon's virtual assistants could be activated for recording by ambient noises that aremistaken for the trigger word. For instance, two employees were interviewed by the software company, Bloomberg, and they reflected that "they heard what they believe was a sexual assault". They claimed that Amazon "reportedly told them it was not the company's job to interfere" (Hern, 2019a).

Moreover, Amazon removed its arbitration clause to allow users to sue the company for collecting voice recordings improperly. Then, they provided the option to delete the voice recordings for users (Cherkassky, n.d.).

Google

Likewise, after a leak of some of its Dutch language recordings, Google admitted that its contractors have access to the voice recordings made by the voice assistants and claimed that they gave the access to recordings to "better understand language patterns and accents, and notes that recordings may be used by the company in its user terms" (Paul, 2019). Similar to the above companies' situation, a spokesman for Google told the Wired Magazine that "only 0.2% of all recordings are accessed by contractors for transcription, and that the audio files are stripped of

identifying user information (Simonite, 2019). A journalist with Belgian public broadcaster VRT, Tim Verheyden, gained access to the recordings from a Google contractor that works as audio file reviewer. This contractor revealed that he "transcribed a recording in which a woman sounded like she was in distress", and he "felt that physical violence was involved" (Simonite, 2019). He also mentioned that they were not provided any clear guidelines on what to do in such cases. VRT produced a video report with the files they gained access to, and there were "recordings of users that had identifiable information, including their address and other personal information, like a family discussing their grandchildren by name, another user discussing their love life, and one user talking about how quickly a child was growing" (Paul, 2019).

Later, Google suspended the transcriptions of voice assistant clips in the EU after the leak of Dutch audio data (Rana & Sampath, 2019). Then, an email was sent to its users who have used any product with voice recognition feature. The email included an announcement indicating that their voice recordings are not being saved, along with a link that allows users to opt-in to the program again. They also stated that, if users chose to opt-in, their voice recordings would be reviewed by staff and related parties (Smith, 2019).

Other Actions

Analyzing the above companies' attempt to resolve the privacy issue, there are several actions they could take that might help with the problem. For example, they could "offer all users the ability to easily and permanently opt out of the data collection practices — one click to say that their voice recording and private information will go nowhere, and will never be seen" (Perez, 2019). They could also follow the EU's GDPR to better protect the users.

For the security issues involving "voice ID", companies could take certain cybersecurity precautions such as two-factor decryption so that the hackers would not be able to control voice

assistants or access users' “voice ID” (Cui et al., 2016). As hackers become better at impersonating people, we want to be able to prevent and stop catastrophic damages caused by voice fraud (Mee & Ozturk, 2020).

Analysis

The cases presented above shows the companies' actions attempting to improve on data protection. It also showed the different concerns that the public have against the companies and voice recognition technology. Sometimes, the concerns were proven to be valid. The ethics principle of justice suggests treating others fairly. It is not fair to collect information on customers without their prior knowledge of the collection. However, companies made effort to treat the users fairly by providing the opt-out option for recording of the voice recognition contained technology. In addition, other actions were provided by the public for companies to reference to when thinking about the next step for data protection. That would also include improving the privacy laws in the United States.

Counter Arguments

According to Walia (2019), privacy controversy over voice assistants misses the point. She made several statements as follows. First, she pointed out that privacy is an illusion. The reason is that almost any form of digital communication technology leaves a trace. Even if the trace is deleted, it would still get tracked, sold, and used to target users. For a long time, people have embraced the free service-for-data tradeoff. They are aware of the cautionary phrase: “once it is on the internet, it is there forever”, and they already forfeit a certain amount of privacy to engage in modern society today.

Second, she claims that there is a huge difference between “hearing” and “listening”. Even though virtual assistants are always listening and waiting for the trigger word, it does not

mean that they are transcribing everything they hear. Moreover, Amazon Echo devices do not have the attention span to secretly eavesdrop on conversations (Gershgorn, 2017). They are "limited from a hardware perspective and incapable of prolonged eavesdropping" (Walia, 2019). Third, virtual assistants need users' data to train themselves and become smarter. Their ultimate goal is to better serve the users, and they could accomplish that by improving their automatic speech recognition (ASR), and their natural language understanding (NLU) capabilities. "Only when ASR and NLU combine forces are our voice assistants capable of truly listening to us, and that is the only thing that makes voice assistants useful" (Walia, n.d.).

As discussed above, the reason voice assistants exist is to assist the public. Companies want advanced technologies to better serve the community. I conducted interviews on students at UVA that own voice recognition products. One student has Amazon's Alexa at home to play music and call family members when needed. Another student uses Apple's Siri to set alarm and check weathers daily. Overall, the interviews' results show that most of the users are satisfied with the products they have. It is not promising to assume the bad in technology when the only side seen is bad. One bad thing should not cover up all the good things it has brought to the society.

Conclusion

Data protection has always been an important topic in this century. People fear what will happen eventually if customer data is not protected properly. To incorporate ethics into this STS topic, customers have the right to be informed of their personal information being stored. Companies and designs should think about the bigger picture and protect the customers.
As the famous quote derived from the poem, Judge Softly, says, "putting yourself in someone else's shoes" (Lathrap, 1895). We know that the recommended actions are easy to say, but hard

to implement, as people usually look at this problem from the user's perspective. We want to be able to understand the companies' perspective, so that we can better analyze the situation.

According to multiple sources like TermsFeed (2020) and Martin (2017), being transparent with the users is significant to effectively communicate with the users. This research identifies actions the companies with voice recognition feature can take to protect their users' personal information and privacy. The Actor Network Theory suggests using relationships between actors to govern a network in order to explain social effects (Crawford, 2020). In this case, we consider the technologies from companies and the public as two actors. They should have interests that guide them to a creation of a new system to produce effects that meet these interests. These interests could be from advancement of voice recognition technologies to better served community.

References

Apple Inc. (2019, August 28). Improving Siri's privacy protections. Apple Newsroom. Retrieved March 30, 2022, from https://www.apple.com/newsroom/2019/08/improving-siris-privacy-protections/

BBC. (2015, February 9). Not in front of the telly: Warning over 'listening' TV. BBC News. Retrieved March 30, 2022, from https://www.bbc.com/news/technology-31296188

Bimber, B. (1990, May 1). Karl Marx and the Three Faces of Technological Determinism. Social Studies of Science, 20(2), 333–351. https://doi.org/10.1177/030631290020002006

Cherkassky, D. (n.d.). The Voice Privacy Problem. Kardome VUI Technology. Retrieved March 31, 2022, from https://www.kardome.com/blog-posts/voice-privacy-concerns

Crawford, T. (2020, September 28) Actor-Network Theory. *Oxford Research Encyclopedia of Literature.* Retrieved 3 May. 2022, from https://oxfordre.com/literature/view/10.1093/acrefore/9780190201098.001.0001/acrefore-9780190201098-e-965.

Cui, H., Paulet, R., Nepal, S., Yi, X., & Mbimbi, B. (2020). Two-factor decryption: A better way to protect data security and privacy. The Computer Journal, 64(4), 550–563. https://doi.org/10.1093/comjnl/bxaa080

Dekate, A., Kulkarni, C., & Killedar, R. (2016). Study of Voice Controlled Personal Assistant Device. International Journal of Computer Trends and Technology, 42(1), 42–46. https://doi.org/10.14445/22312803/ijctt-v42p107

Gershgorn, D. (2017, November 8). The technical reason why Alexa can't listen into your private conversations. Quartz. Retrieved March 31, 2022, from https://qz.com/1121880/the-technical-reason-why-alexa-cant-listen-into-your-private-conversations/

Google. (n.d.). About privacy and personalized ads (formerly known as interest-based ads). Google Ads Help. Retrieved March 29, 2022, from https://support.google.com/google-ads/answer/2549116?hl=en

Hayes, P., & Wagner, J. (2017). Prepare for the voice revolution. Voice assistants. Retrieved March 29, 2022, from https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/voice-assistants.pdf

Hern, A. (2019a, April 11). Amazon staff listen to customers' Alexa Recordings, report says. The Guardian. Retrieved March 31, 2022, from https://www.theguardian.com/technology/2019/apr/11/amazon-staff-listen-to-customers-alexa-recordings-report-says

Hern, A. (2019b, July 26). Apple Contractors 'regularly hear confidential details' on Siri Recordings. The Guardian. Retrieved March 31, 2022, from https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings?CMP=Share_iOSApp_Other

Jones, R. (2018, September 22). Voice recognition: Is it really as secure as it sounds? The Guardian. Retrieved March 29, 2022, from https://www.theguardian.com/money/2018/sep/22/voice-recognition-is-it-really-as-secure-as-it-sounds

Kelly, G. (2019, November 26). Compare the privacy practices of the most popular smart speakers with Virtual assistants. Common Sense Education. Retrieved March 29, 2022, from https://www.commonsense.org/education/articles/compare-the-privacy-practices-of-the-most-popular-smart-speakers-with-virtual-assistants

Lathrap, M. T. (1895). Judge Softly. https://www.aaanativearts.com/walk-mile-in-his-moccasins

Martin, J. A. (2017, April 19). How to keep virtual assistants from sharing your company's secrets. CSO Online. Retrieved October 19, 2021, from https://www.csoonline.com/article/3190837/5-ways-to-keep-virtual-assistants-from-sharing-your-companys-secrets.html

Mee, P., & Ozturk, G. (2020, May 5). Prepare to protect your customers' voices. MIT Sloan Management Review. Retrieved March 31, 2022, from https://sloanreview.mit.edu/article/prepare-to-protect-your-customers-voices/

O'Boyle, B. (2020, September 14). What is Siri and how does Siri work? Pocket-Lint. Retrieved October 19, 2021, from https://www.pocket-lint.com/apps/news/apple/112346-what-is-siri-apple-s-personal-voice-assistant-explained

Olson, C. (2019, April 23). New report tackles tough questions on voice and ai. Microsoft Advertising. Retrieved March 29, 2022, from https://about.ads.microsoft.com/en-us/blog/post/april-2019/new-report-tackles-tough-questions-on-voice-and-ai

Paul, K. (2019, July 11). Google workers can listen to what people say to its AI Home Devices. The Guardian. Retrieved March 31, 2022, from https://www.theguardian.com/technology/2019/jul/11/google-home-assistant-listen-recordings-users-privacy

Perez, S. (2019, April 24). 41% of voice assistant users have concerns about trust and privacy, report finds. TechCrunch. Retrieved March 29, 2022, from https://techcrunch.com/2019/04/24/41-of-voice-assistant-users-have-concerns-about-trust-and-privacy-report-finds/

Ramos, D. (2018, April 16). Voice Assistants: How voice assistants are changing our lives. Retrieved March 29, 2022, from https://www.smartsheet.com/voice-assistants-artificial-intelligence

Rana, A., & Sampath, U. (2019, August 1). Google suspends transcription of voice assistant Clips in Europe. Reuters. Retrieved March 31, 2022, from https://www.reuters.com/article/us-google-european-union/google-suspends-transcription-of-voice-assistant-clips-in-europe-idUSKCN1UR5JX

Sarwar, N. (2021, September 2). Did Apple's Siri assistant violate user privacy? the courts will decide. Retrieved March 29, 2022, from https://screenrant.com/apple-siri-assistant-user-privacy-class-action-lawsuit/

Simonite, T. (2019, July 10). Who's listening when you talk to your google assistant? Wired. Retrieved March 31, 2022, from https://www.wired.com/story/whos-listening-talk-google-assistant/

Smith, D. (2020, August 9). Google's privacy controls on recordings change. what that means for your google home. CNET. Retrieved March 31, 2022, from https://www.cnet.com/home/smart-home/googles-privacy-controls-on-recordings-changes-what-that-means-for-your-google-home/

TermsFeed, & B, R. (2020, February 19). Voice Assistants and Privacy Issues. TermsFeed. Retrieved October 19, 2021, from https://www.termsfeed.com/blog/voice-assistants-privacy-issues/

Waddell, K. (2016, May 24). Why Digital assistants are a privacy nightmare. The Atlantic. Retrieved March 29, 2022, from

https://www.theatlantic.com/technology/archive/2016/05/the-privacy-problem-with-digital-assistants/483950/

Walia, S. (2019, December 27). Why the privacy controversy over voice assistants misses the point - dzone IOT. DZone IoT Zone. Retrieved March 31, 2022, from https://dzone.com/articles/why-the-privacy-controversy-over-voice-assistants

Walia, S. (n.d.). Why the privacy controversy over voice assistants misses the point. RAIN Agency. Retrieved March 31, 2022, from https://rain.agency/raindrops/privacy-controversy-voice-assistants-misses-point