Apple’s Siri: Surveillance in the Modern Technological Age

A Research Paper
in STS4600
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Systems Engineering

By

Amber Ecelbarger

April 30, 2020

Approved by Prof. Kent Wayland, Department of Engineering & Society

Introduction

Voice-controlled intelligent personal assistants (IPAs) have become increasingly popular in recent years as nearly half (47%) of Americans now report using an IPA on their smartphone or in their home (Liao et al., 2019). The accelerated growth of IPAs can be attributed to the scale at which they are being integrated with other technology systems such as home speakers, cellular phones, smart homes, and vehicles. While there are three major players in the mobile voice assistant market, Apple's Siri controls 46%. This makes Apple a clear front runner, as their share is 1.5 times larger than Google's Assistant, which holds the second largest share (Gal, 2018). Siri, which was first launched in 2011, changed the way people interact with their phones by providing a fast, accurate, and hands-free access to information.

As artificial intelligence (AI) technologies continue to develop and become more integrated into existing systems, personal data becomes more important to larger technology companies as data is essential to train AI models and develop personalized user profiles. Large tech companies such as Google, Apple, Facebook, and Amazon have the capabilities to track and profile our behavior over periods of time (*Competition and Data*, n.d.). Collecting excessive user data for analysis is classified as a form of corporate surveillance and is becoming more prominent in the modern technological age. But what concerns does this present for the user? What is really at risk while using these devices? The answer is simple: privacy. When Siri is prompted, your device listens to and stores your data to ultimately benefit Apple. Even when Siri is not in use, the microphone is always partially active as it is expected to respond immediately when prompted with the wake words (*Competition and Data*). My research seeks to understand how and what data Apple collects using Siri and the privacy concerns this presents for its users.

Along with this analysis of Apple and Siri, I will also look into the larger implications around using IPAs as a form of surveillance and how this ultimately impacts consumers. Surveillance, although instinctively categorized as a negative actor on society, has also been viewed in a positive light by a few scholars. In my paper, I will explore both the positive and negative impacts of this surveillance on users. This analysis will provide readers with knowledge to better understand the risks and rewards associated with IPA usage. I will also provide insight as to the role Siri plays in the larger network of corporate surveillance, and whether the company's product and policies are beneficial or harmful to its consumers.

Background and Context

Before exploring privacy concerns with Siri, it is valuable to understand how IPA technology is typically used in today's society. A recent study on Amazon's "Alexa" usage, one of Siri's top competitors, found the most common interactions in U.S. households included checking weather, finding facts, listening to news, controlling other nearby devices, setting reminders or calendar alerts, and playing music (Lopatovska, 2019). In addition to recreational use, IPAs are also used in organizations and other public settings such as hospitals, museums, classrooms, and hotels (Lopatovska). Given the growing ubiquity and rapid expansion of IPAs in public settings, there is a significant amount of research surrounding the privacy and security concerns of these digital assistants. This research informs my analysis and sits at the heart of this paper.

To better understand the role Siri plays in the network of corporate surveillance, it is beneficial to discuss how the voice assistant came to be and how Apple collects data from it. In April 2010, Apple purchased the Siri application prototype from the Sandford Research Institute for over 200 million dollars (Blank, 2019). Following this acquisition, Apple continued to enhance Siri using natural language processing, data collection, and personalization. These features allowed Siri to better understand the user, send data for processing, and provide an appropriate response. In order to process data, Apple uses iCloud, a cloud-based storage system which allows the company to store and process user data off of the physical device (Keller, 2018). Apple revealed in their 2018 iOS Security Guide that the company uses both Amazon and Google's commercial cloud storage systems to store their iCloud data (Fleishman et al., 2018). Given the high value of data, this novel cloud technology allows and incentivizes larger tech companies to collect and store detailed information on each individual user. Though Apple uses cloud-based computing, the company claims when it comes to Siri, the user's Apple ID is not connected to the IPA software and the majority of conversations and requests are processed directly on the individual's device (*Privacy*, n.d.).

Apple's impressive storage capabilities and company claims pose the question: what data does Siri actually collect from its users? To answer this question, I turned to Apple's website which claims, "Siri is designed to protect your information and enable you to choose what you share" (*Ask Siri, Dictation & Privacy*, 2019). The company elaborates on the data collection process and explains that when you make a request to Siri, it sends certain personal data to Apple to help process and respond to the request. This ancillary data includes contacts, music and podcasts you listen to, names of your photo albums, and more. Apple also claims all requests are associated with a random identifier, meaning that these requests are anonymized and cannot be traced back to specific user (*Ask Siri, Dictation & Privacy*, n.d.). Apple updated Siri's data collection methods and privacy policy in October of 2019, after the company faced scrutiny on these practices earlier that year. I will analyze how changes to Apple's policy and data collection methods impact user privacy later in this paper.

Literature Review

Scholarly research on surveillance often highlights either the potential harm or underlying benefits of IPAs as a form of surveillance, but fail to consider the potential of both outcomes. Current literature frequently scrutinizes “corporate players” and larger technology companies, without factoring in each company’s motivations and values when placing a moral judgement on them. Emily West, an esteemed scholar at the University of Massachusetts, is an exception to this general research oversight (West, 2019). West looks at Amazon holistically and argues the company’s surveillance is a positive service for its consumers. This analysis aligns with what I intend to do, but West does not balance her assessment by considering the negative impacts of the company’s actions. Researching this area, I also found that many of the sources covering the controversy with Apple’s data collection through Siri focused heavily on the practice itself and less on what data was actually collected and whether or not this truly presents a privacy concern for users. My research intends to close this gap in understanding and analysis by looking at both the positive and negative impacts of surveillance on users in the context of Apple, the data they collect, and their unique role within this corporate network.

Theoretical Framework

To better understand the role IPA technology plays in society, I will consider how IPAs and society are mutually shaped by one another. This perspective suggests that society and technology are not mutually exclusive, but are rather constantly influencing and shaping one another. To apply this framework, I will study how IPAs have evolved and changed to meet consumer needs and in turn, how consumers are shaped by their use of this technology. Apple’s revisions to their privacy policies will reveal how Siri evolved to meet the desires and concerns of users. The alleged harm and benefits associated with IPAs will demonstrate how Siri and other

IPAs shape and impact consumers. I will also explore how further integration could change the way society operates in the future to understand how the technology and its consumers might continue to shape one another. The societal aspects I discuss in this paper span from personal freedoms of privacy and security, to widespread public services such as law enforcement, and healthcare.

Research Question and Methods

My research question is how has Apple's Siri evolved with user privacy concerns and what does this suggest on a larger scale about IPAs and surveillance? I will investigate how Apple's Siri first created and then evolved with its users' privacy concerns to understand how this type of relationship can be generalized to IPAs and the issue of surveillance. My overall approach to this question involves a literature review and content analysis of relevant sources to pinpoint reoccurring themes. The first part of my research seeks to understand how the use of Siri has raised user concerns around privacy. To do this, I looked at Apple's website and other articles to evaluate the company's current and past mission statement, values, and policy surrounding privacy to determine if any glaring changes have been made. I also referenced several news articles where Apple's privacy protections were put into question and perform a content analysis on these sources. To determine how the attitudes on privacy influenced the development of Siri, I looked at Apple's response to these concerns.

The second part of my research question explores the role of IPAs in surveillance. To perform this analysis, I collected scholarly work that discusses surveillance in the modern technological age and more specifically, the role of IPAs in this surveillance. I found several sources that discuss either the positive or negative aspects of surveillance for users. Using this evidence along with my knowledge of Apple, I then commented on where I believe Apple fits in on this surveillance spectrum. I ensure my data collection process is rigorous by analyzing sources with a level of skepticism that drives me to further research the author or host cite to ensure the data and information can be validated.

Apple Inc.

In Apple's latest mission statement, it promises the "best user experience to its customers through its innovative hardware, software, and services" (MSA, 2019). In 2011, Apple's CEO, Tim Cook, shared the company's seven core values, which embody Apple's attitude and priorities prior to Siri's release. These values reflected the goals set forth by Steve Jobs and focused on innovation rather than privacy, but even as of 2016, privacy concerns were not as widely recognized as they are today. Reviewing Apple's website, as of 2020, we can see that privacy is now listed as one of their core values, demonstrating a shift in priorities and marketing strategy within the company (*Privacy*, n.d.). The change in company values sheds light on the prominence of privacy concerns around IPA applications and other technology brought about in recent years. As engineers are becoming aware of these user concerns, changes are being made to technology, policies, and data collection practices. These reactionary changes to the product are evidence that society is shaping IPA technology, even at a company as influential as Apple.

Privacy Concerns with Siri

Although the addition of "privacy" as a core value shows that Apple is aware of its users concerns, it does not mean this new value is being embedded into the technology. It is essential to investigate whether Siri actually protects the privacy of its users and maintains this company wide value. On their website, Apple clearly recognizes protecting personal privacy as a fundamental goal of its operations; however, if this is such an important issue to the company, why has it received so much scrutiny on this topic? According to a news report by The Guardian, Apple failed to disclose that a small portion of Siri recordings are screened by private contractors around the world, leading users to question their privacy with the device (Hern, 2019). According to Apple, this surveillance is used to improve Siri and dictation; the data collected is used to help the software better understand the user and what is said (Hern). Apple attempted to reassure the public that these soundbites were anonymized: the user's Apple ID is never attached to the recording. But Apple's lack of transparency around their use of human oversight in their data analysis process proved incredibly concerning to many of its users.

Following this scrutiny, Apple vowed to make necessary changes to their data collection process with Siri. In the current Privacy Statement on their website, the company highlights the these changes which include no longer collecting audio recordings of Siri interactions and allowing users to opt in or out of the data collection process used to improve Siri and dictation (*Privacy - Features*, n.d.). Apple's response to the societal expectations of privacy is one example of how IPA technology is mutually shaped by society. The changes made to Apple's Privacy Statement ultimately impacted and shaped Apple's voice assistant, proving that the technology did in fact evolve with to meet society's values and concerns. I will detail these changes later in this paper when discussing how companies attempt to mitigate potential harms of IPA surveillance.

IPAs as Surveillance

Now that we have reviewed Siri's data collection methods and privacy policy, I will now look into the larger implications of IPAs acting as a form of surveillance in society. To understand these implications, I will investigate how IPAs facilitate surveillance in the modern technological age and how consumers view these devices. IPAs have proven to be extremely beneficial to consumers as they provide quick access to information and can complete user tasks by leveraging artificial intelligence, machine learning, and natural language processing. The privacy concerns these devices present is largely a result of their use of cloud computing and the massive amount of consumer data transmitted.

An article published by several scholars including Jason Pridmore, an associate in the Department of Media and Communication at Erasmus University in the Netherlands and Michael Zimmer, a privacy scholar and associate professor at Marquette University, explores the intercultural differences in perspective on surveillance in households. The article details the results from a study which sought to understand the cross-cultural differences that play a role in understanding IPAs as surveillance found interesting results. This study, which used several focus groups from the US and the Netherlands, found that while some Americans expressed ambivalence towards the privacy concerns with IPAs, many Americans felt their interactions with the devices were mundane and of little interest to the company who may be collecting their data (Pridmore et al., 2019). While Dutch participants expressed more apprehension towards IPA data collection, American participants generally accepted that data collection and profiling was an inevitable result of using IPAs and many were unbothered by these consequences in exchange for the benefits they receive such as convenience and tailored advertising (Pridmore et al.).

The findings of this study suggest that some consumers do not have concerns about the privacy of IPAs, but this prompts the question of whether this lack of concern comes from users not caring about the effects of this surveillance, or if they are naïve to the consequences and implications it brings about. Thinking on a larger scale, the continual development of this technological surveillance could potentially allow for the misuse of personal data by private companies and the government. Constant surveillance may not seem like such a bad thing on a granular level and protections such as encryption and legislation make it more difficult for this corruption to occur, but in many cases, users hastily grant companies the right to collect their data without considering the impact of this in unexpected areas. For example, insurance companies could adjust rates based on the data and knowledge unknowingly collected about you (*Hey Siri, Am I Being Watched?*, n.d.). If an IPA device or some other technology were installed in your vehicle, automobile insurance companies may adjust rates based on the data collected on your driving patterns, and similarly, health insurance companies may adjust rates based on the food you purchase (*Hey Siri, Am I Being Watched?*, n.d.). The continual integration of smart devices into your homes, automobiles, and other devices enable a vast network of surveillance with potentially harmful effects. However, this is not to say that the consequences of surveillance are all negative; they may result in beneficial outcomes, such as lower insurance rates.

Surveillance can result in both positive and negative outcomes for the consumer, but can we analyze whether a company's surveillance is more beneficial than harmful? If we consider the data collection motives of Apple and other popular IPA developers such as Amazon and Google, we can see how different motives may drive the level and type of surveillance a company conducts. The same article published by Pridmore, Zimmer and other scholars argued that Amazon's motivation for data collection revolves heavily on their desire to drive personalized sales on their website, while Google uses data to build and improve its targeted advertising platform (Pridmore et al.). The article argued that compared to its competitors, Apple was viewed more positively by participants who attested the company appears more focused on selling hardware rather than user data (Pridmore et al.). The perception of these companies is often reflective of their intention and how they use data and should be considered when exploring the positive and negative impact of surveillance on users.

Potential Harms of Surveillance

IPA devices have been widely criticized by users who fear their security and privacy are at risk. Two potential harms of this surveillance include the ability of law enforcement to request IPA recordings as evidence and the potential for hackers to control your devices or access your information.

Could the data stored on your device be used as evidence against you? This alarming ethical question has been tested in recent years given law enforcement's request of IPA data in a 2015 murder investigation. During, State vs. Bates, a murder trial filed in the Arkansas Circuit Court, law enforcement issued warrants for several digital devices present at the scene which included an Echo and Alexa, two of Amazon's IPAs (*The Privacy Implications of Virtual Personal Assistants*, 2018). Amazon responded with an attempt to squash the subpoena fearing the release of this data would result in severe distrust from its users. Amazon argued the conversations between users and their devices should be subject to protection under the First Amendment. The controversial questions of this case were never fully answered as Bates, the defendant, eventually agreed to have the recordings released. Although not addressed by the court, this case also brings into question protections under the Fourth Amendment Third Party Doctrine, which essentially states that "people are not entitled to an expectation of privacy in information they voluntarily provide to third parties" such as an internet or cellular service provider (Ii, 2014, p. 2). This doctrine essentially grants the government the right to a massive amount of information on an individual such as the websites they visit, emails they send, and more. The data these devices collect and the right of legal officers to obtain this data is an unsettling reality one should recognize when using an IPA device.

Because a single IPA device amasses a significant amount of personal data, it is important to consider how secure this information actually is on these devices. Apple, when compared to its tech counterparts, is better at protecting the privacy of its users. As mentioned previously, in the case where Apple does process or store data associated with Siri interactions on its servers, the data is tagged with a random identifier and is thus, not linked to the user (*Privacy*, n.d.). On the other hand, Google and Amazon both store records of the user's queries so all historical conversations with their IPA can be accessed on the user's account (Vlahos, 2019). A case study using Amazon's Alexa and Google's Home sought to determine whether these devices employ the necessary security mechanisms to protect users from attacks and found three shortcomings: weak single-factor authentication, no physical presence-based access control, as well as insecure access control on Alexa-enabled devices (Lei et al., 2018). The paper also discusses practical attacks that result from the indicated vulnerabilities which could result in scenarios where a hacker is able to abuse Alexa's capabilities to place a fake order or even a home burglary if Alexa opens a door on command.

Along with these internal vulnerabilities, IPAs have also shown to be externally prone to hacking. A study from researchers at the University of Michigan and Japan's University of Electro-Communications found that Siri, Alexa, and other IPAs could be commanded by shining a laser at them (Metz, 2019). This could be done from hundreds of feet away and anyone with the knowledge and resources could command the device to perform a task such as opening your garage door. The threats identified in this section demonstrate the potential negative impact IPAs have on users and the ways in which Siri is equipped to combat these harms.

Underlying Benefits of Surveillance

Along with these potential harms, I will discuss some of the prospective benefits of this surveillance on society. The two I will highlight in this paper are smart homes and increased access to health services.

While previous sections detailed home security threats associated with IPAs, I will elaborate on the potential benefits afforded by IPA's in the home assuming the device is secure. Emily West, an associate professor in the Department of Communication at the University of Massachusetts, details the positive impact Amazon's IPA has on its customers. West argues that through the use of Alexa, the company is essentially selling home surveillance as a service to its users in that consumers can monitor and protect their personal spaces through the interconnected web of technology (West, 2019). Using IPAs and other technology in the home as surveillance would not only deter criminals from entry, but would also allow homeowners to save energy by allowing them to easily monitor and control their heating, cooling, and electricity usage (*How Surveillance Systems Will Impact Our Lives In the Next 10 Years?*, 2017).

In the healthcare industry, this technological surveillance could mean close monitoring of patients with certain conditions and quicker access to medical aid when needed. According to a guide which outlines national patient safety goals, a key metric in assessing patient safety is how adequately a risk is identified and mitigated (Dzou, 2019). Identifying and mitigating patient risk is essential as failing to do so poses a threat to other faculty members and patients. Surveillance technology which combines a camera with the AI capabilities of an IPA, would help hospital staff to better monitor situations before they become larger threats as the camera can identify visual cues and call attention to potential risky situations and the system would also be responsive to vocalized patient requests. Similar to how they would operate in a hospital, IPAs could also be used in the home to care for elderly loved ones or even sick children. So, while technological surveillance is often viewed in a negative light and presents several concerns to consumers, it also has the potential to beneficially shape society's healthcare operations and access to security.

Where Does Apple Fit In?

When we consider Apple's use and operations of Siri, it is important to consider both the positive and negative impacts of IPA surveillance on society and Apple's role as a large corporate player in the tech industry. Apple claims their products diverge from other players as they are designed to keep your personal information private and secure (Brandom, 2019). My research affirms the security of their data collection methods and adequacy of the policy revisions around Siri. Given this evidence, I would argue that Apple has sufficiently dealt with past concerns of privacy and improved their operations to merit them recognition as a company that upholds the values of privacy and security.

Does Apple's surveillance create potential harm for its users? If we consider the company's motive for data collection discussed previously, Apple is more focused on creating high quality hardware than collecting user data. Because the company is more focused on their product than gleaning value from their data, the privacy threat of data collection is reduced for consumers. Apple's policy revisions and updates to their data collection process also reveal the company lacks motive to amass an enormous amount of data on each of its consumer. Given this evidence, I would argue that compared to its competitors, Apple makes it easier to protect users against some of the harmful impacts of surveillance. As for the beneficial impacts of surveillance that I noted, Apple has already ventured into the smart home industry with Apple TV and Home Pod. While the company does not currently play a large role in the healthcare industry, this presents a major opportunity in the future.

One thing we cannot ignore is that these companies are all competing for a larger customer base and market share. Apple, Amazon, Google, and other big tech companies may have different business models, data-collection methods, and motivations, yet all are driven to constantly improve the intelligence of their IPAs and services. Given data is necessary to train AI models and thus, improve IPA software, one could argue the data-collection efforts of larger tech companies will likely increase in the future, adding capabilities to an already massive network of corporate surveillance. While my research shows Apple is less concerned with the collection of user data and combats many of the potential harms associated with surveillance, with the ever-changing technological age and corporate race, Apple should expect a certain level of surveillance from its own users to ensure its values and asserted protections are adequately upheld.

References

Ask Siri, Dictation & Privacy. (n.d.). Apple Support. Retrieved February 7, 2020, from https://support.apple.com/en-us/HT210657

Blank, A. (2019). The Real Siri: Past, Present, and Future. | CCTP-607: “Big Ideas”: AI to the Cloud. https://blogs.commons.georgetown.edu/cctp-607-spring2019/2019/05/05/the-real-siri-past-present-and-future/

Brandom, R. (2019, March 26). Apple wants to be the only tech company you trust. *The Verge.* https://www.theverge.com/2019/3/26/18282158/apple-services-privacy-credit-card-tv-data-sharing

Competition and Data. (n.d.). Privacy International. Retrieved February 1, 2020, from http://privacyinternational.org/explainer/2293/competition-and-data

Dzou, C. (2019, March 26). Using Surveillance Technology to Improve Patient Safety and Care. *Verkada.* https://www.verkada.com/blog/using-surveillance-technology-improve-patient-safety-care/

Fleishman, G., Contributor, S., May 22, M. |, & PDT, 2018 5:00 am. (2018, May 22). *How to find out where Apple stores your iCloud data (spoiler: You can’t exactly)*. Macworld. https://www.macworld.com/article/3274584/where-does-apple-stores-your-icloud-data.html

Gal, P. B., Shayanne. (n.d.). Siri owns 46% of the mobile voice assistant market—One and half times Google Assistant’s share of the market. *Business Insider.* Retrieved February 1, 2020, from https://www.businessinsider.com/siri-google-assistant-voice-market-share-charts-2018-6

Hern, A. (2019, July 26). Apple contractors "regularly hear confidential details" on Siri recordings. *The Guardian*. https://www.theguardian.com/technology/2019/jul/26/apple-contractors-regularly-hear-confidential-details-on-siri-recordings

Hey Siri, Am I Being Watched? The Future of Privacy with the Internet of Things. (n.d.). Viterbi Conversations In Ethics. Retrieved March 3, 2020, from https://vce.usc.edu/hey-siri-am-i-being-watched-the-future-of-privacy-with-the-internet-of-things/

How Surveillance Systems Will Impact Our Lives In the Next 10 Years? (2017, December 7). CHRO. https://www.chro.org/surveillance-systems-will-impact-lives-next-10-years/

Ii, R. M. T. (2014, June 5). *The Fourth Amendment Third-Party Doctrine*. (CRS Report No. R43586). Retrieved from Congressional Research Service website: https://fas.org/sgp/crs/misc/R43586.pdf

Keller, J. (2018, June 1). iCloud: Everything you need to know! *IMore*. https://www.imore.com/icloud-everything-you-need-know

Lei, X., Tu, G.-H., Liu, A., Li, C., & Xie, T. (2018). *The Insecurity of Home Digital Voice Assistants—Vulnerabilities, Attacks and Countermeasures*. 1–9. 2018 IEEE Conference on Communications and Network Security (CNS), Beijing. https://doi.org/10.1109/CNS.2018.8433167

Liao, Y., Vitak, J., Kumar, P., Zimmer, M., & Kritikos, K. (2019). Understanding the Role of Privacy and Trust in Intelligent Personal Assistant Adoption. In N. G. Taylor, C. Christian-Lamb, M. H. Martin, & B. Nardi (Eds.), *Information in Contemporary Society* (pp. 102–113). Springer International Publishing. https://doi.org/10.1007/978-3-030-15742-5_9

Metz, R. (2019, November 5). Researchers used a laser to hack Alexa and other voice assistants. *CNN*. https://www.cnn.com/2019/11/04/tech/alexa-siri-laser-attack-research/index.html

MSA. (2019, May 23). Apple Mission Statement 2020 | Apple Mission & Vision Analysis. *Mission Statement Academy*. https://mission-statement.com/apple/

Pridmore, J., Zimmer, M., Vitak, J., Mols, A., Trottier, D., Kumar, P. C., & Liao, Y. (2019). Intelligent Personal Assistants and the Intercultural Negotiations of Dataveillance in Platformed Households. *Surveillance & Society*, *17*(1/2), 125–131. https://doi.org/10.24908/ss.v17i1/2.12936

Privacy. (n.d.). Apple. Retrieved March 6, 2020, from https://www.apple.com/privacy/

Privacy—Features. (n.d.). Apple. Retrieved April 3, 2020, from https://www.apple.com/privacy/features/

The Privacy Implications of Virtual Personal Assistants. (2018, August 7). *BusinessWest*. https://businesswest.com/blog/the-privacy-implications-of-virtual-personal-assistants/

Vlahos, J. (2019, March 26). Smart talking: Are our devices threatening our privacy? *The Guardian*. https://www.theguardian.com/technology/2019/mar/26/smart-talking-are-our-devices-threatening-our-privacy

West, E. (2019). Amazon: Surveillance as a Service. *Surveillance & Society*, *17*(1/2), 27–33. https://doi.org/10.24908/ss.v17i1/2.13008