Intersect, Vol 13, No 1 (2019)

I Think I Can Trust Alexa, But How Much?

Jiten Jwalant Bhatt
University of Virginia

Introduction

As our society becomes increasingly data-driven, it becomes progressively imperative that we strive to keep our data safe and in hands we trust. Software and businesses collect and utilize vast amounts of data to operate. Much of this data is user data, which has produced controversy regarding who can collect it and who owns it. Combined with the risk of security breaches, these rights may be seen as a slippery slope that leads to misuse or detriment to the users, ranging from unwanted targeted advertisements or spam being sent through private contact methods to financial loss or physical harm.

Most people find it difficult to function in today's society without smartphones, given the ubiquity of social media. Social media is an example of a technology for which people have generally accepted that some of their data may be visible to the developers or possibly the public due to enough confidence in its security, lack of foresight to be concerned, or perhaps obliviousness to the consequences of a data breach. Regardless of the reason, a significant number of people use this technology enough to adopt it. I consider newer technologies such as artificial intelligence (AI) personal assistants, which can behave most effectively by learning first about humans, or more specifically, the individual(s) each one serves; in other words, they require user data to perform well. The term "AI personal assistant" is vague, as it is debatable whether these assistants should even be considered artificial intelligence; for the purposes of this exploration, I define an "AI personal assistant" as a software agent that can accomplish a variety of tasks requested by a user, usually via multiple types of inputs such as voice commands and stored user data.

In my STS thesis, I aim to explore which factors are causing people to adopt artificial intelligence personal assistants and Internet of Things (IoT) devices despite the risks, in order to assess what actions different actors are taking or should be taking. I will consider existing research and sociotechnical perspectives on AI personal assistants; this is because AI personal assistants are still a growing technology with an expanding user base, so the adoption process is still under way. Studies and surveys on usage of AI personal assistants may provide insight on what may cause people to entrust their data knowing that it could be stolen in a security breach or misused by the entity to which it is provided. Furthermore, AI personal assistants are a precursor to IoT, and studying them may provide insight on how trust in IoT will affect its adoption. To learn more about the adoption of these technologies, I aim to understand how AI personal assistants are used, how much personal data people are comfortable providing to them, and how people feel about the risk and reward of using them. However, as AI personal assistants are currently in the adoption process, they cannot give us the full picture. It is also worth considering similarities to the adoption process of some currently ubiquitous technologies such as social media. I also delve into the factor of trust between different actors to guide my discussion of how the adoption process is shaped, and what needs to be done to ensure a favorable outcome for the users, government, and developers.

The problem of understanding the public view of AI personal assistants and their adoption process is complex, due to the large amount of people, organizations, and other actors involved. Actor-Network Theory (ANT), as proposed by Bruno Latour, suggests that we analyze a problem or topic by viewing it as a shifting network of relationships between actors, which are anything or anyone involved. I will apply ANT to the case of AI personal assistants; I can then analyze the relationships between the different components of the developing company as well as actors they interact with, such as the developers and managers, the users, and attackers or adversaries. ANT will allow me to consider the difficulties of dealing with trust in technology from the perspectives of each of the actors and see how they contribute to the resulting trust or distrust by the people.

A variety of factors have been found to influence an individual's propensity to trust a technology. In one study of user trust in technology, the authors mention that other similar studies name reliability, usability, design, and social presence as major factors. However, they also note that user-specific traits such as age and "inherent trust" may also be significant factors (Xu et al., 2014). Inherent trust may be difficult to measure; however, it does seem plausible that some individuals are more trusting than others. Nevertheless, trust is certainly not synonymous with adoption. For example, social media is one technology that is widely used despite varying levels of trust by its user base. According to a study on user confidence in social networks, 74% of users believe their data is not safe on social networks (Tiganoaia et al., 2017).

Working towards the goal of trust

What exactly is meant by trust? Between people, it is the confidence that one person will do as they say and not make a mistake or go back on their word. Trust is a little different when it is between a human and a machine; while it might seem a bit less likely that the device will betray the user, there is also a level of confidence that the user has in the device's capabilities. For example, if I asked Amazon Alexa to order more napkins online, I might be concerned about Alexa's ability to understand me, to find the right kind of napkins, and to properly place the order with the right payment options, address, etc. While Alexa might not directly choose to betray a user's trust or fail due to limited abilities, a user might not trust the developers and, therefore, fear that their data is at risk of being shared or leaked to unauthorized parties.

The complexity increases even further for IoT devices. One paper lists several types of trust, all of which are relevant for an IoT system and may influence how a user views it. Data perception trust, which consists of how devices receive information from the outside world, has more variability than AI personal assistants, which primarily rely on microphone input. There are more points of vulnerability, which increases the difficulty of data transmission and communication trust as well as system security and robustness (Yan, Zhang, & Vasilakos, 2014). For the purpose of simplifying the actor-network for my discussion, I enumerate these relationships as a few general categories of trust in this context. There is trust between devices and services in a system, between users and developers, and between users and devices. While trust may not be a necessary condition for adoption, as demonstrated by social media, establishing user trust in a product and/or the developers is an important step that contributes to adoption. In many cases, trust must be earned, especially when a user's data is being made potentially vulnerable.

Even if the developer has implemented a secure system that keeps user data safe and has no intentions of misusing it, they need to convince the users of this. One method is the implementation of user interface and user experience (UI/UX) design that gives the user the feeling that their data is being processed securely (Hochleitner, Graf, Unger, & Tscheligi, 2012). For example, when entering credit card information online, a user might feel safer if the webpage looks more professional or shows words such as "secure" or "protected." Another way to help users feel secure is providing them with privacy-protecting technologies. An example of this is a built-in mechanical webcam cover on a laptop, which is used to prevent unwanted webcam access in the event of an attacker attempting to remotely enable it. Analogous techniques may be used for AI personal assistants and IoT devices, such as mechanical switches to disable recording or network access.

The developer also has to keep in mind that because the goal of the UI/UX design is the promotion of trust in addition to usability, they have to treat the user-software interactions as a relationship where the user should not feel betrayed for any reason. This may involve mitigating misplaced animosity in the event that a related component behaves in an unfavorable manner. For example, if a user tries to use an Alexa app such as Spotify, but the app has received an update and isn't working properly, the user may blame Alexa rather than the app's developer, considering Alexa unreliable.

These are some ways to improve user-device trust, however they are difficult to quantify. Even if the user feels that their data has been processed securely because of the UI/UX, their trust could be easily lost due to news of a security breach. It is equally important that their data is actually processed securely; this can be especially difficult to ensure in an IoT network, where any one of several devices could be a point of entry for an attacker.

Why would I use Alexa?

AI personal assistants interpret and draw conclusions from large amounts of data in order to improve, but this means that they would be hindered by a lack of real user data from which to learn. Furthermore, even after development, personal assistants keep track of user data to improve their daily performance and convenience for the users. But researchers question whether there are other reasons data is store; it is also difficult to know exactly how much of it is stored and if it's being collected by the company who sold the personal assistant.

In addition to having difficulty trusting a company producing AI personal assistants due to their motives, users should also be concerned about security flaws in the developer's system that can be exploited by malicious third parties. One paper on trusting personal assistants highlights many major vulnerabilities, including wiretapping a system, compromising it like any other computing device via any form of hacking, impersonating a user to provide malicious voice commands, or unwanted sound recording (Chung et al., 2017). These vulnerabilities can even surface by accident; one woman reported that Alexa recorded a conversation and sent it to someone in her contacts. Amazon reported that this occurred due to an unlikely series of events in which Alexa misinterpreted noise as commands to record and then email ("An Amazon Echo recorded a family's conversation, then sent it to a random person in their contacts, report says," n.d.). Alexa may or may not always be recording, but because it is always listening, the possibility for these cases is nontrivial. Amazon issued a statement of their intent to work to further reduce the likelihood of accidents like this, but even if these vulnerabilities are supposedly patched, how can the user base be assured of the safety of their data?

Additionally, from a non-technical perspective, it's not surprising that AI personal assistants seem risky. Some people view AI personal assistants not just as a step forward in making tasks--such as playing music, or making a to-do list--easier, but also a step toward bringing technology closer to our level (Shulevitz). In her article, Shulevitz argues that switching from manual computer inputs to voice inputs is not just "a matter of switching out the body parts used to accomplish those tasks" but a change in societal status for the personal assistants. The Social Construct of Technology (SCOT) suggests that human actions shape the development of technology. What we consider to be good AI personal assistants are those that understand us as users and respond in a natural, understandable way; in other words, more human-like features generally improve the user experience. After all, a natural conversation is much easier than to put in the effort to restructure a command you want to make into broken language that a less intelligent AI can understand. However, this idea where AI personal assistants rise to the same level as humans is yet another fear that could prevent or hinder the adoption of personal assistants.

Why do others use Alexa?

Given the security risks, apprehension surrounding AI, and mistrust of data collection and unwanted recording, why do 47% of Americans report using AI personal assistants on either their phones or at home (Liao, et al.)? While AI assistants in general are viewed with skepticism, it's important to consider the actors involved in individual cases, particularly those most successful. Consider Alexa, Amazon's AI personal assistant. Based on data from interviews with college students and home-owners over the age of 40, the consensus was that while there is some risk involved, Alexa is convenient and useful. College students preferred to keep Alexa outside of the bedroom, making it less likely for private phone calls and other sensitive conversations to be heard. (R. Rustagi, personal communication, March 4, 2019). Some of the older interviewees were aware of the possibility that Alexa is always listening but they would not consistently take action to avoid having sensitive conversations near the device. In other cases, people are simply not aware of the risks or consequences of privacy and security breaches, or just trust the developers and the security of the system. There is also a degree of trust in the law to keep the developers in check.

The primary trust-earning factor for Alexa in particular is the developer's reputation. Amazon Alexa is a feature implemented in several Echo devices, which are generally considered smart speakers. Over the past two years, Amazon Prime membership has nearly doubled to over 80 million. Amazon and Google have increased their market share in smart speakers in the last few quarters as well ("Amazon Increases Global Smart Speaker Sales Share in Q4 2018, While Google's Rise Narrows the Gap and Apple Declines"). While Alexa and Prime are different products, Prime membership suggests an increased tendency to purchase Amazon products as well as customer confidence in Amazon, indicating a nontrivial correlation with Amazon Echo sales. Alexa is also available as a free mobile application and integrates with a number of smart devices such as smart plugs and lights. Even if someone does not have an Echo product, there are other devices with which users can utilize Alexa, increasing the overall presence of Alexa beyond the primary market of smart speakers. In contrast, consider how Facebook's reputation affects usage of its new products. Facebook's Portal device is essentially a video-calling device with a camera, microphone, and screen. However, due to recent events that raised issues with trust in Facebook as a company, it was initially very poorly received, as articulated by one of many reviews:

> But the bigger issue most people will have with the Portal is that it's an always-watching and always-listening device *connected to Facebook*. The device's release was reportedly delayed for several months in the wake of the Cambridge Analytica scandal, in which Facebook was pilloried for failing to put strict controls on data shared with third-party developers. And just as Facebook prepared to release the device, the company revealed that a new data breach had compromised the accounts of more than 50 million people. (Seifert, 2018)

Many widespread technologies in today's society utilize vast amounts of data, but concerns arise regarding both the security of this data and the people's trust in the technologies themselves. However, despite clear vulnerabilities in AI personal assistants that also apply to the IoT, AI personal assistant usage continues to grow. This is similar to how social media is so widely used and influential despite several scandals involving user data breaches and sales. This suggests that trust is not the only factor that contributes to adoption. The social presence of the technology and the people's perception of and relationship with the companies involved also play a role. Amazon prides itself on its customer-centric approach, which includes providing reliable services and making things right with its customers when it errs, in order to not lose customers to competitors. While not directly related to the success of only one of its products–Alexa–it does improve likelihood of Amazon customers trusting and purchasing Amazon-branded products.

With both rapidly increasing Prime membership and Echo product sales numbers, customers are becoming more likely to give Amazon-branded products a chance. Even if Alexa seems like a data risk to a customer, the customer might be inclined to use it anyways since it is an Amazon product. These strong relationships allow even some questionable actions to go unnoticed or be condoned by the consumers. For example, Amazon recently continued the process of licensing its facial recognition software to government and law enforcement agencies despite employee protests (Statt, 2018). AWS CEO Andrew Jassy notes that Rekognition has done a lot of good, such as prevent human trafficking, and that it would be detrimental to take this technology away. However, he acknowledges the risks and reassures us that the terms and services protect proper usage, suggesting that misuse will result in a ban. He also notes that "it's the role in the responsibility of the government to help specify what the guidelines of regulations should be about technology." This seems like a way to push the blame towards government regulators and claim that anything bad that happens is only because the regulators allow it, but Jassy does have a point. Some users are well-aware of data privacy issues and make efforts to limit risk of unwanted dissemination of private information such as avoiding discussion of sensitive information near microphones. However, most people aren't aware that they need to take precautions at all, and to expect them to all be informed and careful is unrealistic.

Developers of technology and the government have a responsibility to ensure safety for users. But concerns arise when these are the people who cannot be easily trusted, which could lead to dystopian results. User data is valuable, and entities that possess it stand to profit from it. Even if a developer is entirely benevolent and has no plans to exploit its users, they can still suffer from irresponsible software system and security architecture. This issue has substantial implications for IoT as well; while AI personal assistants have vulnerabilities such as wiretapping or an attacker that uses your device to order shipments, a compromised IoT system could lock someone out of their own house or even cause physical damage to the house or house owner. AI personal assistants and IoT devices that collect user data increase the scope of these vulnerabilities, so it will become increasingly important to guarantee user data privacy and protection.

The General Data Protection Regulation, recently passed in the European Union, is the first major regulation that requires companies to notify consumers of data breaches, collect data legally while maintaining records on how and why they do so, and protect user privacy in a broad sense. While the GDPR only directly applies to the EU, it is pushing companies to change their strategies on how they handle data protection and privacy. Unfortunately, in the USA, there is not yet an equivalent regulation. There are some smaller laws related to privacy, but they are easily circumvented due to being limited to certain sectors; some examples include HIPAA, COPPA, and FERPA, which protect healthcare, children, and student information, respectively. Punishments for most infractions are rarely more than small fines, which are especially insignificant when the culprits are companies earning billions in revenue. The GDPR, in comparison to existing legislation, can issue fines of up to 4% of the culprit organization's revenue. The Obama Administration tried to push for a "Consumer Privacy Bill of Rights" which shares several ideas with the GDPR, however despite further iterations over the last few years, federal legislation protecting private information has not yet been passed ("We Can't Wait," 2012). The GDPR makes no specific mention of AI personal assistants, however this may not be necessary for effective regulation. Given the primary weakness of existing legislation being sector-specific, federal legislation should aim to be as broad as possible, protecting people's data collected by more than just technology companies, enforcing a culture of user privacy in all relevant industries.

GDPR only went into effect last May, so it's difficult to draw concrete conclusions on its strengths and weaknesses; nonetheless, it has caused some notable events such as a $50 million fine to Google for insufficiently informing user on how data is used for advertisement personalization (Fox, 2019), as well as a personal data request made to Amazon that resulted in 1700 voice recordings not belonging to the requester being incorrectly sent ("Amazon Customer Receives 1,700 Audio Files Of A Stranger Who Used Alexa," 2018). While $50 million is not difficult for Google to pay, it's also well below the maximum 4% fine that could have been given. The fact that GDPR is being enforced and utilized by both lawmaking bodies and users, as well as the increased likelihood that the US will follow suit with federal legislation could be the push needed for companies to invest in improving their data privacy and security practices.

Conclusion

While a significant amount of users may boycott a product as protest against data misuse or insecure systems, according to a paper focused on surveying human trust in an IoT context, "the human heuristic handling of risks, threats and opportunities is not without its faults, but use of trusted proxy devices and the trust we have in recognized brands and companies will enable us to trust many services without too much hesitation" (Køien, 2011).

In addition, we need to keep in mind that companies can amass users of their product or service without explicitly establishing trust, and so precautions need to be taken to ensure that this behavior is not abused. These precautions are not just important for reducing future issues with AI personal assistants and social media, but for IoT, which has potential to be far more widespread and a much larger market. The limitations of my paper may include an inability to address factors such as large-scale change in the perspectives of people. It is difficult to extrapolate results of discussion of AI personal assistants and IoT to other products and services; different technologies have different advantages, disadvantages, and concerns that could lead to trust issues.

Regardless, in this age where handling of user data is so ubiquitous and vulnerable, developers, companies, and law enforcement must all take steps to protect consumers. Developers may strive to build secure systems, but when companies stand to profit from selling user data and do not suffer from legal repercussions, there is a very dangerous conflict of interest. The government needs to pass legislation similar to the GPDR, with design that maximizes scope across industries to reduce loopholes caused by a focus on sector-specific legislation. Of course, more focused legislation on top of GDPR-like regulations are ideal. It's possible that federal regulation may already be sufficient for some technology, but AI personal assistants and IoT pose complex challenges given their dependence on user data to operate and will likely require more scrutiny.

While many users do not have confidence that their data is safe when collected and utilized by AI personal assistants and IoT systems, most will continue to use them due to the platforms' reliability and usefulness. Therefore, it is imperative that lawmaking bodies push for legislation to follow the lead of GDPR, to drive a culture of data privacy via a top-down approach; with laws pressuring companies to prioritize security and discourage data misuse, developers will be incentivized to build systems that meet these requirements. Complete trust by users may not be essential and likely isn’t possible, but for sustainable adoption, a vulnerable user base must be protected so they can continue to use new technology such as personal assistants with confidence that they are not put at risk.

References

Amazon Customer Receives 1,700 Audio Files Of A Stranger Who Used Alexa. (2018, December 20). Retrieved December 17, 2019, from NPR.org website: https://www.npr.org/2018/12/20/678631013/amazon-customer-receives-1-700-audio-files-of-a-stranger-who-used-alexa

Amazon Increases Global Smart Speaker Sales Share in Q4 2018, While Google's Rise Narrows the Gap and Apple Declines. (2019, February 20). Retrieved February 27, 2019, from https://voicebot.ai/2019/02/20/amazon-increases-global-smart-speaker-sales-share-in-q4-2018-while-googles-rise-narrows-the-gap-and-apple-declines/

An Amazon Echo recorded a family's conversation, then sent it to a random person in their contacts, report says. (n.d.). Retrieved May 2, 2019, from Washington Post website: https://www.washingtonpost.com/news/the-switch/wp/2018/05/24/an-amazon-echo-recorded-a-familys-conversation-then-sent-it-to-a-random-person-in-their-contacts-report-says/

Chung, H., Iorga, M., Voas, J., & Lee, S. (2017). Alexa, Can I Trust You? *Computer*, *50*(9), 100–104. https://doi.org/10.1109/MC.2017.3571053

Fox, C. (2019, January 21). Google hit with £44m GDPR fine. *BBC News*. Retrieved from https://www.bbc.com/news/technology-46944696

Hochleitner, C., Graf, C., Unger, D., & Tscheligi, M. (2012). *Making Devices Trustworthy: Security and Trust Feedback in the Internet of Things*. 6.

Køien, G. M. (2011). Reflections on Trust in Devices: An Informal Survey of Human Trust in an Internet-of-Things Context. *Wireless Personal Communications*, *61*(3), 495–510. https://doi.org/10.1007/s11277-011-0386-4

Lee, J.-H., & Song, C.-H. (2013). Effects of trust and perceived risk on user acceptance of a new technology service. *Social Behavior and Personality: An International Journal*, *41*(4), 587–597. https://doi.org/10.2224/sbp.2013.41.4.587

Montague, E. N. H., Winchester, W. W., & Kleiner, B. M. (2010). Trust in medical technology by patients and healthcare providers in obstetric work systems. *Behaviour & Information Technology*, *29*(5), 541–554. https://doi.org/10.1080/01449291003752914

Statt, N. (2018, November 8). Amazon told employees it would continue to sell facial recognition software to law enforcement. Retrieved February 20, 2019, from https://www.theverge.com/2018/11/8/18077292/amazon-rekognition-jeff-bezos-andrew-jassy-facial-recognition-ice-rights-violations

Strategy Analytics: 2018 Global Smart Speaker Sales Reached 86.2 Million Units on Back of Record Q4 | Strategy Analytics Online Newsroom. (n.d.). Retrieved February 27, 2019, from https://news.strategyanalytics.com/press-release/devices/strategy-analytics-2018-global-smart-speaker-sales-reached-862-million-units

Szymczak, H., Kücükbalaban, P., Lemanski, S., Knuth, D., & Schmidt, S. (2016). Trusting Facebook in Crisis Situations: The Role of General Use and General Trust Toward Facebook. *CyberPsychology, Behavior & Social Networking*, *19*(1), 23–27. https://doi.org/10.1089/cyber.2015.0450

Tiganoaia, B., Cernian, A., & Niculescu, A. (2017). The risks in the social networks #8212; An exploratory study. In *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (Vol. 2, pp. 974–977). https://doi.org/10.1109/IDAACS.2017.8095232

University of Minnesota, Bapna, R., Gupta, A., University of Minnesota, Rice, S., Texas A&M University, … New York University. (2017). Trust and the Strength of Ties in Online Social Networks: An Exploratory Field Experiment. MIS Quarterly, 41(1), 115–130. https://doi.org/10.25300/MISQ/2017/41.1.06

We Can’t Wait: Obama Administration Unveils Blueprint for a “Privacy Bill of Rights” to Protect Consumers Online. (2012, February 23). Retrieved December 17, 2019, from Whitehouse.gov website: https://obamawhitehouse.archives.gov/the-press-office/2012/02/23/we-can-t-wait-obama-administration-unveils-blueprint-privacy-bill-rights

Xu, J., Le, K., Deitermann, A., & Montague, E. (2014). How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology. *Applied Ergonomics*, *45*(6), 1495–1503. https://doi.org/10.1016/j.apergo.2014.04.012

Yan, Z., Zhang, P., & Vasilakos, A. V. (2014). A survey on trust management for Internet of

Things. *Journal of Network and Computer Applications*, *42*, 120–134. https://doi.org/10.1016/j.jnca.2014.01.014

Yao, Y., Viswanath, B., Cryan, J., Zheng, H., & Zhao, B. Y. (2017). Automated Crowdturfing

Attacks and Defenses in Online Review Systems. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1143–1158). New York, NY, USA: ACM. https://doi.org/10.1145/3133956.3133990