How loyal is your Alexa? Imagining a Respectful Smart Assistant

William Seymour
University of Oxford
Oxford, OX1 3QD, UK
william.seymour@cs.ox.ac.uk

Abstract
Smart assistants are the current must-have device in the home. Currently available products do little to respect the autonomy and privacy of end users, but it doesn't have to be this way. My research explores a speculative 'respectful' assistant which is more socially aware, and treats its users in a more nuanced way than occurs at present. Mixing computer science, philosophy, and art, the project uses a combination of user studies and technical comparison to discover a potential future for the smart digital assistant.

Author Keywords
Internet of Things; Smart Homes; Digital Assistants; Respect

ACM Classification Keywords
Human-centered computing [Empirical studies in interaction design]; Human-centered computing [Natural language interfaces]; Human-centered computing [Human-centered computing Sound-based input / output]; Security and privacy [Social aspects of security and privacy]

CHI'18 Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada
ACM 978-1-4503-5621-3/18/04.
https://doi.org/10.1145/3170427.3180289

Introduction
Smart home assistants such as Amazon Echo and Google Home are the must-have gadget of the moment. They are convenient because they allow us to perform a variety of tasks that would normally require physical effort (even if only on a smartphone) using only verbal commands. They also allow us to leverage the functionality of many different Internet of Things (IoT) devices which aid in automating our homes.

We have a tendency to conceptualise these devices as both intelligent and subservient, but often such beliefs do not match reality. Users lack control over the software that runs on their devices, are offered a Hobson's choice[1] over which types of personal data are collected and transmitted back to the manufacturer, and must rely on themselves to draw information out of their assistants rather than be alerted to pertinent information.

[1] Characterised as 'take it or leave it'.

In this paper I describe my ongoing research, which is a speculative exploration into a different kind of smart personal assistant: one which is considerate and loyal to its users in a way that current offerings are not. I will also explore how a device prioritising its end user over its manufacturer can influence the trust which users place in the technology they buy for their homes by creating a prototype assistant. By observing user interactions with the prototype and comparing the privacy guarantees it offers to the current market leaders I hope to gain an insight into what a probable smart assistant of the future might look like, and what path the current state of affairs might have to take in order to get there.

This paper begins by exploring the deficiencies in current smart assistants, before briefly covering the prior work that the project builds on. Finally, I describe the laboratory experiments and user studies that I plan to undertake in order to understand the speculative future that the respectful personal assistant inhabits.

The Problems with Current Smart Assistants

One common feature of the smart assistants of today is the lack of choice available to the user. Beginning from when the user first open the box and configures their device, through to when the device is disposed of, if a user wishes to use the device in any capacity they are obliged to adhere to the terms and usage patterns proscribed by the device manufacturer.

Control Over Data

Having purchased a smart assistant, the first step of the setup process involves the user agreeing to a lengthy set of terms and conditions concerning (amongst other things) the data that the manufacturer of the device is able to collect, process, and store about them. This process is more of an ultimatum than an agreement—refusal to accept often requires the user to return the device for a refund.

After this, the device begins to silently collect data and metadata about the user which are then sent back to the manufacturer. Not only is it likely that the user has not read the end user license agreement (EULA), and thus does not understand what data are being collected, but this information is mostly transmitted out of the device encrypted in a way that prevents the user (or any other parties) from auditing it.

Unlike in other areas of life, users are not presented with the option to customise the services they receive so as to have them better conform to their own privacy preferences. Ideally users should be able to opt out of services which require permissions they do not wish to grant. Similar to the Hobson's choice imposed by EULAs, this lack of choice forces consumers to choose the 'least bad' option when considering smart assistants, instead of the best fit for them.

Additionally, given the privileged and central network location that smart assistants often occupy (especially when acting as a hub for simpler IoT devices), they are perfectly placed to detect leakage of personal data from *other* devices before it leaves the local network and escapes forever onto the internet. This is, however, not a feature offered by any of the popular choices in the smart assistant market.

R.E.S.P.E.C.T[2]

There is something with the above paragraphs which intuitively feels 'wrong' in a way in which other forms of data collection do not. We allow smart assistants into the most intimate spaces in our lives: our homes, our bedrooms,

[2] A pun on the chorus of the 1967 song by Aretha Franklin

and our bathrooms. Much like any other guest permitted into these spaces, there are certain standards that devices placed there are unconsciously held to. I believe that these standards can be conceptualised in a single term, respect.

But what are the implications if we position respect as an important goal when designing smart assistants? One could successfully argue that the current iterations of assistants **do** already respect the laws and regulations of the countries they are sold in. But while adhering to clearly defined boundaries is necessary in order to be respectful, it is not sufficient.

The other type of respect embodied by the current generation of devices can be summed up as *obstacle respect* [2]. This is when an agent acknowledges the wants and needs of another in order to better serve its own goals. Agents which show obstacle respect need only do as little as is required in order to keep the support of those who's cooperation they require.

So what additional types of respect woukd an ideal smart assistant display? We often take respect to signify a deeper view of an agent over and above the surface-level resources we desire to extract from them. In order for a subject to show this *recognition respect* there must be an acceptance of the object as having intrinsic value. Our smart assistants must appreciate us as being more than merely an efficient way to extract money and data that can be mined for profit.

The logical next step after recognition respect is *care respect*, where one sees the object of one's respect as something that should be cared for and protected. The very properties that originally engendered the respect and subsequent respectful behaviour cause the subject to want to take care of the object. Unlike the results of obstacle respect, care respect manifests itself in a way that values the long term health and needs of the agent or object being respected.

These behaviours are all ones which would be expected of another human sharing our personal spaces, and as smart assistants increasingly try and become more life-like they need to make more of an effort to follow suit. In the end this may also be beneficial for the device itself. Much like a friend who upon entering our house refuses to help with anything that doesn't directly benefit themselves, we are less likely to trust disrespectful smart devices with our data.

Methods of Interaction
Interactions with current smart assistants follow very well defined patterns. Wake words prompt the device to listen, and then there are often a small number of predetermined commands which can be used to interrogate the device for data, or to cause it to effect various actions around the home (such as turning lights on and off, or raising the heating temperature). Communication with smart assistants is a one-way process, with verbal interactions initiated by the user in a fashion not dissimilar to using an Application Programming Interface (API).

This model of communication follows decades of science fiction in which superheroes and star-ship captains query highly intelligent computers. But there is a whole host of pertinent information that a home assistant is privy to that does not function well under the currently used 'pull' notification paradigm. This can take on trivial guises, such as new emails or instant messages, but could also be unknown devices connected to the home network, or suspicious packet traces from existing devices. Should be expect our assistants to act more like Data and less like HAL[3]?

[3] Intelligent computers from *2001: A Space Odyssey* and *Star Trek* respectively

Similarly to how showing respect can help to build trust between a device and its users, interacting in a more human manner can help us to feel easier about having an assistant in our homes. So long as the device does not overstate its capabilities, this can also make it easier for users to state complex requests in a concise and natural way [3].

Background and Related Work

Speculative Design

In order to fully judge the efficacy of the proposed solutions to the problems with smart assistants, the project will use speculative design methods in order to gather authentic reactions and feedback to possible evolutions of the smart assistant. *Speculative design* describes design methods which can help to both imagine the future and critique current work, and are explored by James Auger [1].

These techniques exist alongside similar ones such as design probes and design fictions, but differ in that they connect the design objects more strongly to everyday life, as well as the contemporary items from which they were born. By keeping speculations close to the here and now, speculative design avoids straying too close to science fiction and thus dislocating objects from what people consider to be real. Similar techniques (design fictions) have been used by Lindley et al. to explore privacy in the context of smart homes and data protection regulation, a topic closely related to this research [5].

Related Work

The nature of this project means that it crosses disciplinary boundaries and touches different areas of computer science, psychology, and art. What follows is a brief overview of the main topics that the project builds on: the dangers of anthropomorphising software, machine-led deception, respect, and data privacy in the context of smart 'things'.

When addressing the lack of human warmth encountered by users interacting with smart assistants, one must take care not to overstate the capabilities of the software. In their 1992 paper, Friedman and Khan argue that as assistants and similar pieces of software are able to appear increasingly intelligent, there is a danger that users will associate greater functionality and capability to the system than it actually possesses [3]. While there are certainly situations where this problem can be avoided, one can imagine this manifesting itself in a number of ways, such as a user expecting their assistant to know that reminding them 30 minutes before a meeting was OK but 30 minutes before a flight was not (for example).

According to Andreas Matthias, promoting (even unintentionally) this disconnect between real and perceived capabilities can be thought of as deception [6]. This can rob users of autonomy when they rely on their assistant to do something it is incapable of, causing diminished trust in the device when the disconnect is discovered. Moral justification for deceiving users will vary from domain to domain, but the willingness of the user to be deceived should always play a part.

When considering deceptive smart assistants it is also important to remember that deception forms a normal part of human interaction. In order for an assistant to communicate in a fully 'human' way it would need to be able to tell low consequence lies in appropriate circumstances. Van Kleek et al. also consider the ways in which software can facilitate socially deceptive actions of users [4], and David Traum amongst others explores the issues around creating virtual personas that are not always truthful [8].

A large amount of philosophical thinking on respect has been collated and succinctly summarised by the Stanford Encyclopedia of Philosophy [2].

There has been much discussion in academic publications and general media about the data privacy problems presented by badly written or legal but undesirable IoT firmware and associated smartphone applications. Both Shklovski et al. and Van Kleek et al. have written about the data sharing practices of apps, including how user perceptions change when they become more informed about the data that their devices are sharing about them [7, 9].

Research Approach

The first phase of the project has been to analyse the existing literature concerning respect, human interactions, and deception. Through this a number of design goals have become apparent that will be taken forward into the development of a speculative smart assistant.

The Smart Assistant

The construction of the respectful smart assistant itself will form phase two of the project. Based on an Intel NUC kit augmented with a speaker and microphone, the unit will be of a similar size and form factor to currently available smart assistants. The majority of the functionality of the unit will be provided through Jasper[4], an open source platform for developing voice-controlled applications.

Building a new system, rather than modifying an existing solution such as the Amazon Echo, allows for more control over how modules are handled (e.g. the use of passive modules), as well as a selection of different speech-to-text and text-to-speech engines. This introduction of choice means the prototype will be able to give stronger privacy guarantees than conventional assistants.

[4] https://jasperproject.github.io

Software Modules (i.e. 'Skills')

The respectful assistant will be capable of interacting with a limited set of smart home devices in order to help users accept the speculative artefact as real. Examples of what will be possible include adjusting home lighting and retrieving data from scales and security cameras.

The assistant will be able to notify users when suspicious network activity is detected (such as transmission of personal data in clear text). Not only will this add to user agency with respect to understanding which web services they use are insecure, but it will also be able to alert them to insecure (or malicious) devices on the home network which are exfiltrating personal information. Users will then be able to block or restrict the network access available to these devices.

In keeping with the overarching theme of respecting the user, the assistant will attempt to detect undesired practices by merchants such as differential pricing, and other situations where content shown to users is manipulated as a result of web tracking.

The respectful assistant will also be able to identify new devices connected to the home WiFi network, alerting the user when devices are added unexpectedly (such as when the user is out of the house).

User Studies

The third phase of the project involves inviting participants into the lab to interact with the smart assistant. The use of a controlled environment allows for initial reactions to be captured, and for a greater understanding of how users adapt their behaviour to the features and limitations of the assistant.

It is important to be able to link the sentiments of participants with observed and documented behaviour during this freeform part of the study. The information gathered will be used to adapt and improve the assistant software in order to make it more design faction than design fiction.

Finally, participants will be invited to take the assistant home with them for a short period of time, in order to investigate how a respectful system might be used on a day-to-day basis.

Comparative Experiments
In the lab a number of experiments will be run to capture the technical differences between the respectful assistants and current solutions with high market shares. This will help build an understanding of how a future respectful assistant might fit into the consumer market, as well as how the current state of affairs might transition into a more respectful one.

Conclusion

Smart assistants will continue to play an increasingly impactful role in our lives, with relatively little research done into the social effects of voice activated assistants. Through the use of speculative design principles this research aims to realise a more socially aware smart assistant whilst at the same time exploring respect as a design goal for smart devices. User studies and comparative technical analysis will be used to support these theoretic goals with empirical evidence.

REFERENCES

1. James Auger. 2013. Speculative design: crafting the speculation, In Digital Creativity. *Digital Creativity* 24, 1 (2013), 11–35.
2. Robin S. Dillon. 2016. Respect. In *The Stanford Encyclopedia of Philosophy* (winter 2016 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
3. Batya Friedman and Peter H. Kahn Jr. 1992. Human Agency and Responsible Computing: Implications for Computer and System Design. In *Journal of Systems Software*.
4. Max Van Kleek, Dave Murray-Rust, Amy Guy, Kieron O'Hara, and Nigel Shadbolt. 2016. Computationally Mediated Pro-Social Deception. In *CHI 2016*.
5. Joseph Lindley, Paul Coulton, Haider Ali Akmal, and Bran Knowles. 2017. Anticipating GDPR in Smart Homes Through Fictional Conversational Objects.
6. Andreas Matthias. 2015. Robot Lies in Health Care: When Is Deception Morally Permissible?, In Kennedy Institute of Ethics Journal. *Kennedy Institute of Ethics Journal* 25, 2 (2015), 169–162.
7. Irina Shklovski, Scott D Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. 2014. Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use. In *CHI'14*. 2347–2356.
8. David Traum. 2012. Non-Cooperative and Deceptive Virtual Agents.
9. Max Van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J Weitzner, and Nigel Shadbolt. 2017. Better the devil you know: Exposing the data sharing practices of smartphone apps. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 5208–5220.