Opening Statement
26th November 2019
Dr Benjamin R Cowan
Assistant Professor, School of Information & Communication Studies, University College Dublin
Member of Science Foundation Ireland's ADAPT Centre

First of all I'd like to thank the Chairperson, members of the committee, parliamentarians and all who have been involved in calling this committee. I feel honoured to be have been invited to give evidence.

I'd firstly like to give the committee an overview of the research I lead in digital assistants at UCD and at ADAPT. I lead research into the user experience and user behaviour with voice-based devices, investigating what the barriers and concerns of users are, as well as opportunities there are in improving our experience with these devices. Our recent focus in particular in both UCD and ADAPT is in the growing importance of user trust with these devices, as well as other artificial intelligence technologies more broadly. Colleagues such as Dr Marguerite Barry also work specifically on the role of data ethics around these interactions. I'd like to specifically thank Dr Barry and Gianluigi Riva who have contributed significantly to this evidence.

Prevalence: It is hard to under-estimate the prevalence and increasing reach of voice-based technologies. Smart speakers have been a catalyst for a huge growth in the popularity of voice as a way of interacting with our technology. It is estimated that the number of smart speakers installed will grow to 207.9 million by the end of this year, with a lot of this growth being from China and the US. The data estimates that this will grow up to 500 million units by the end of 2023. For Ireland, just under 10% of households now own one of these devices. Note that this is not taking into consideration that if we own a smartphone we all already have one of these agents in our pockets at all times. The reach of this technology is huge. It's also clear that the places where we use these are growing. Voice assistants are being used in the car, in healthcare contexts as well as in the home. Especially with smart speakers, they are becoming social devices by default, being placed in public spaces in the home, with multiple people interacting with each device. Friends, relatives, parents and children, people who live there as well as visitors. They are the gateway to the Internet of Things, whereby we use commands to control devices in our house such as lights, alarms, doors and other devices. This is the context within which these technologies are being used, and in which data is gathered.

How these work: A number of these devices use wake words, using the microphone of the device to listen in the background, activating when these are recognised and uttered. The "intelligence" is almost entirely on the cloud, where the audio is analysed, speech recognised, action is planned and delivered to the agent application. Most of the time, none of this is on the device. Location and other forms of data can also be used to personalise your experience, like giving you weather for your location or adapting the

information you get based on local norms. The user voice data in the form of recordings or transcription of the recognised utterance as well as the system's action/response are commonly available through a user profile. These logs are clear with either the audio recorded, what it thought you said to the system and the response the system gave, with the user being given the opportunity to delete these audio files if they wish in a number of cases.

What is the data currently used for? The data gathered is thought to be used to improve the way the system operates. The more data these organisations have, the better they can get the system to be. For instance, this can be useful to feed the AI techniques used to develop these systems, helping to improve recognition, its understanding of the language used, along with improving its responses to queries. These techniques benefit from large amounts of data, so the use of user data is a big advantage when trying to make and improve the experience of these systems.

Voice data issues: There is no two ways about it, these devices record your speech. Your voice recordings can give people the information you send to these devices of course, but it is other signals that can be equally troubling. Paralinguistic cues allow accurate estimates of age, sex and even native language. It can also be used to build a version of your voice for particular commands to impersonate you. Getting rid of a password is easy, but getting a new voice if other devices use voice authentication and someone has copied it is much harder. With the use of third-party applications, this data is also likely being transmitted, shared or stored, using their infrastructure, which may potentially have security flaws. As anyone who has used these devices knows, these can also record you unintentionally, picking up and storing this audio. Add to the mix recent news around this voice data being listened to by human transcribers, and it is clear that this is an issue that needs to be addressed.

So what do we need to focus on when we think about data for digital assistants:

1) Being clear about what "always on" means: These devices mean that, in effect, there is a microphone in every home/phone that is constantly on, maybe waiting for a particular word or utterance, but it is listening. These can, at best, record accidentally, but at worst could be intercepted and used to monitor users. This seems unnecessarily intrusive to me. It may not be to others, but users have to be made aware of this.

2) Clarity on why data is being stored, who accesses it and what it's combined with: Currently the reason that data is kept is opaque to the user, summarised by being used for "improvements to the system". Data gatherers must be more explicit about how this data is used in terms of said improvements, tracking, profiling and sharing across an organisation as well as with third party organisations. This needs to be explicit to the user, along with how the data is paired with other streams to influence the experience.

3) Giving back user control: Currently the user has no control over who can access this data, or what it can be used for. This means that there is no opportunity for users to have an active ongoing voice in how it is used. It also sets the competitive advantage of the big data players, who can use their oceans of data to improve their system with competition being left with little data to play with. This makes it hard for smaller start-ups who may change the way these systems operate to compete. Giving the user's control of their data would allow them to choose where and with whom their data recordings resides. It may also allow a boom in research in the area as users could potentially donate their data to be used for non-profit research if they desire. Crucially this would put the user in control.

4) Consent for all users: Currently these systems are used in public spaces by multiple users. Think of a smart speaker in a kitchen or a living room. Audio is being captured of a number of different users, neighbours, visiting relatives, children. None of these have consented to their data being recorded and stored but all may be being recorded. We need to discuss new consent mechanisms for these devices.

5) Discussing on how to design for privacy: Consideration needs to be taken of how we can include privacy into digital assistant design as standard. Some smart speakers have the option to turn off the microphone so it is not "always on". Push to interact mechanisms would also reduce the likelihood of recording accidentally. But we also need to be aware of what this means for the user. This type of design decision has a trade-off of convenience for the user, when they may want to use these in hands busy, eyes busy situations. These types of options on the device go some way to reducing the surveillance potential. Ensuring that audio is stored and processed on the device rather than sent to the Cloud would also go some way to reducing the risks from hacking.

Conclusion: Our work shows that privacy is indeed a concern for users. Although it may not seem to influence user behaviours yet, it is in the company's and the governments interests to address this head on. The data we are talking about here is not a set of clicks, a search history or a set of cookies, it is our voices, which is perceived as something far more personal. A hack or misuse of this data would be significant and the threat of this is real. We should, as users, have our eyes open as to what it means when we invite a digital assistant into our home.

References & resources:

https://voicebot.ai/2019/04/15/smart-speaker-installed-base-to-surpass-200-million-in-2019-grow-to-500-million-in-2023-canalys/

[https://voicebot.ai/2019/10/11/over-20-of-uk-households-have-smart-speakers-while-germany-passes-10-and-ireland-approaches-that-milestone/]

Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F. (2019). Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *26*(3), 17.

Alepis, E., & Patsakis, C. (2017). Monkey says, monkey does: security and privacy on voice assistants. IEEE Access, 5, 17841-17851.

Nautsch, A., Jiménez, A., Treiber, A., Kolberg, J., Jasserand, C., Kindt, E., ... & Abdelraheem, M. A. (2019). Preserving Privacy in Speaker and Speech Characterisation. Computer Speech & Language.

Chung, H., Iorga, M., Voas, J., & Lee, S. (2017). Alexa, can I trust you?. Computer, 50(9), 100-104.