

CDEI Snapshot Series

Smart Speakers and Voice Assistants

September 2019

**Centre for
Data Ethics
and Innovation**

About this CDEI Snapshot Paper

The Centre for Data Ethics and Innovation (CDEI) is an advisory body set up by the UK government and led by an independent board of experts. It is tasked with identifying the measures we need to take to maximise the benefits of AI and data-driven technology for our society and economy. The CDEI has a unique mandate to advise government on these issues, drawing on expertise and perspectives from across society.

The CDEI Snapshots are a series of briefing papers that aim to improve public understanding of topical issues related to the development and deployment of AI. These papers are intended to separate fact from fiction, clarify what is known and unknown, and suggest areas for further investigation.

To develop this Snapshot Paper, we undertook a review of academic and grey literature, and spoke with the following experts:

- **Dr Hamed Haddadi**, Imperial College London
- **Dr Christopher Burr**, University of Oxford
- **Mukul Devichand**, BBC
- **William Tunstall-Pedoe**, AI Entrepreneur
- **Emma Higham**, Google
- **Prof Verena Rieser**, Heriot Watt University

Contents

Summary	2
Introduction and context	3
What are smart speakers and voice assistants?	4
Data collection, use, and privacy	8
The disruptive effect	13
Ideas for further action	19
Conclusion	21

Summary

Reading time: 15–20 minutes

Smart speakers have soared in popularity, with expectations that the number of households using them will grow significantly over the coming year. This developing technology undoubtedly presents exciting opportunities, including making it easier for people to access the online world or control other devices. But public concerns about smart speakers have been expressed. Many of these focus on the seemingly intrusive aspects of the devices and the use of the data captured. Others have raised questions about their longer-term disruptive impact on the consumption of information, user profiling and people's relationship with technology.

This paper focuses on three areas which pull together the key topics of public concern:

1. **Voice recordings are routinely processed by both machine learning models and human review** to improve capabilities of products and drive further innovation. Some such recordings are made when devices wrongly detect the wake-word. Public awareness of this data capture and use appears low and could be better explained to users.
2. Recordings collected by voice assistants provide platforms with new troves of data which they may potentially use to **profile customers** in new ways – such as analysis of sentiment or even aspects of mental health. The extent to which this occurs is opaque. Yet the **types of insights about users that can be generated from this data are only likely to increase**. While such applications could be beneficial, the analysis of voice data to make inferences about individuals must be conducted in a transparent and accountable manner.
3. Smart speakers present new and engaging ways for people to consume material online. However, a shift away from screen-based content may leave **the major platforms in potentially powerful positions, with their devices becoming key gateways to the web**. The longer-term impact of this is debatable but regulators need to be aware of it as an emerging issue.

1. Introduction and context

The use of smart speakers, or voice assistants, like Alexa, Siri, Cortana and Google Assistant has soared with at least one in 5 homes in the UK estimated to be using them.¹ Priya Abani, a Director at Amazon, has said “we basically envision a world where Alexa is everywhere”.² Smart speakers are increasingly prevalent in people’s homes, scanning for a wake-word before responding to a user’s instructions. Part of their attraction has undoubtedly been their novelty and apparent flexibility. From telling jokes to simply laughing on request, people have enjoyed interacting with the devices which have human-like personalities. As they become embedded in more homes, voice-enabled devices look set to fundamentally change how people engage with technology. Freeing people’s hands while they issue verbal instructions, the disruptive element of this nascent technology is exciting. But for many, it is also unnerving.

Although users increasingly rely on them as part of their daily routines, the devices’ constant monitoring of people’s most intimate environments can seem intrusive. Media articles around the world have raised concerns about the devices supposedly listening to private conversations and sharing recordings. The assistants record snippets of audio from homes, potentially enabling companies to gather whole new types of data for analysis. While data from people’s online interactions has long-been collected, the analysis of audio data is new.

Concerns about the more disruptive impact of smart speakers tend to be hypothetical and therefore harder to address. The technology may transform how users engage with the Web. While this brings opportunities it potentially increases the power of the major technology platforms still further, enabling them to wield more influence over people’s online interactions.

With concerns about smart speakers regularly generating public attention, an exploration of the relevant issues is particularly timely.

1 OFCOM (2019) *Nations and Regions Technology Tracker*.

2 Pierce, D (1st August 2018) *Inside the lab where Amazon’s Alexa takes over the world*, wired.com

2. What are smart speakers and voice assistants?

This snapshot focuses on voice-enabled assistants – internet-connected software which responds to voice commands to provide content and services, interacting with users via digitally-generated voice responses.

At the time of writing there are four major voice assistants available in the UK: Amazon's Alexa, Microsoft's Cortana, Google Assistant, and Apple's Siri. Competition among the manufacturers is intense, with low budget options now available. Beyond these types of smart home devices, similar technology is increasingly featured across a range of connected devices including smartphones, tablets, PCs and even children's toys.

Given the limits of this short paper, we primarily look at the impact of voice assistants through their use in smart speakers.

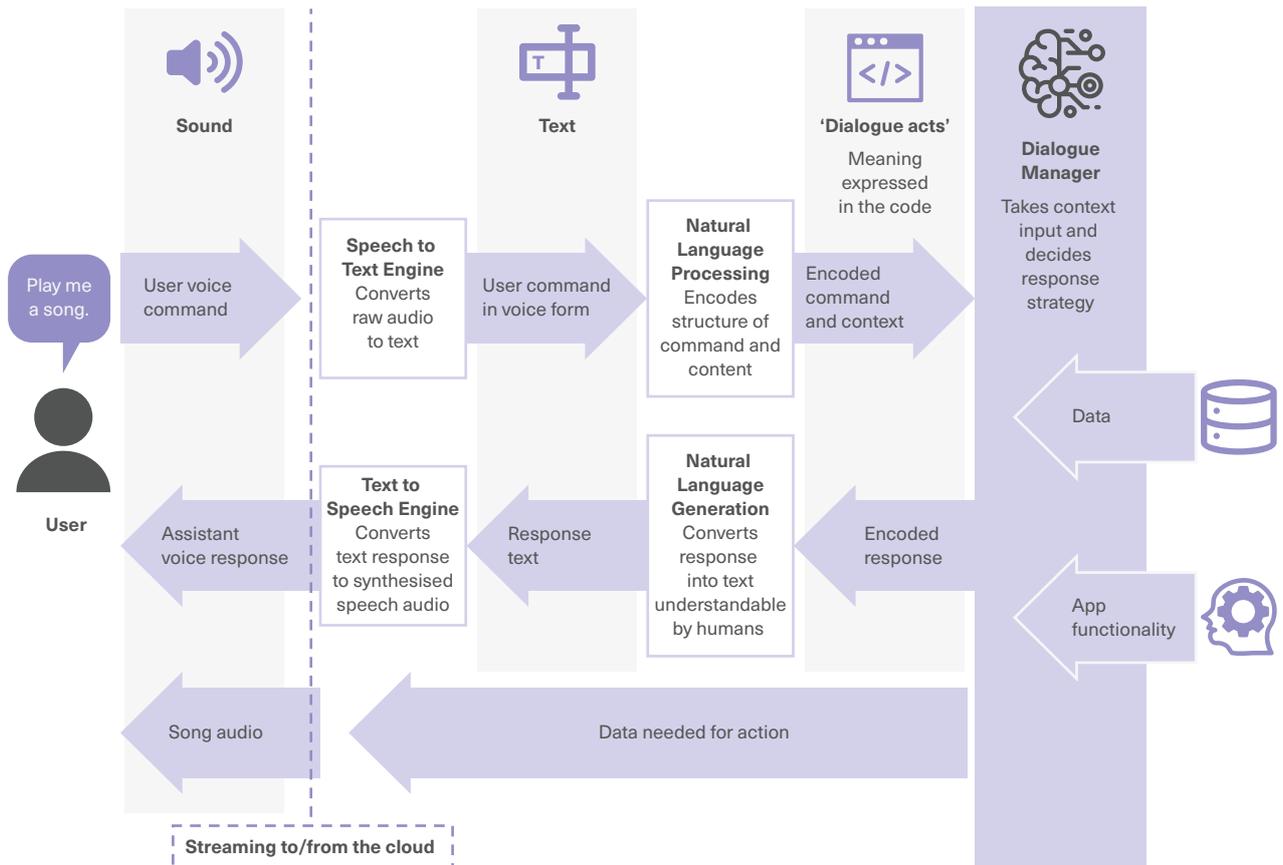
How do they work?

Despite the diversity of products available, the way smart speakers work is broadly similar. At any given moment, the device holds a continuous 'buffer' of the last few seconds of sound recorded from its environment, which it scans for the wake-word. Only once the wake-word is detected does the device begin recording and streaming audio to the cloud for analysis and storage.³

In the cloud, the audio is converted using speech-to-text technology, and run through a form of natural language processing that turns it into a machine-understandable structure of meaning. The results of this are then passed to a 'dialogue manager' that determines the best response to the interaction – for example, playing music, or running a search query. In doing so, it will take into account the perceived intent of the user, additional context such as the device's location, and will also decide whether to execute a response like streaming music via a specific app.

Once an appropriate response has been determined, the voice response heard by the user is produced. The audio is streamed back to the device, along with any further actions needed to meet the user's request.

³ Although Amazon is reportedly developing functionality to detect the wake word at any point in a sentence. See *Amazon Alexa: 'Pre-wakeword' patent application suggests plans to process more of your speech*, The Register (5 June 2019).



Adapted from diagram by Prof Verena Rieser, Heriot Watt University⁴

How are people using them?

Those who own smart speakers generally embed them into their daily routines and call on them to support relatively basic tasks like providing entertainment and retrieving information. A report by the Reuters Institute for the Study of Journalism⁵ found that playing music is the feature used most by the vast majority of users in the UK (84%). Answering general questions, providing weather updates and setting alarms were the next most popular uses. The same report notes that the more whimsical features such as 'tell me a joke', or playing general knowledge games were also appreciated.

4 Rieser, V. and Lemon, O. (2011) *Reinforcement Learning for Adaptive Dialogue Systems* (pg 12), Springer Publishing.

5 Newman, N. (November 2018) *The Future of Voice and the Implications for News*, Digital News Project, Reuters Institute for the Study of Journalism, University of Oxford.

Benefits

From setting timers while cooking, to calling friends without pressing any buttons, the devices enable multitasking and support people's busy lifestyles. But for those who are not confident using technology or have a disability that makes using a keyboard or screen interface challenging, the technology may be life-changing. The devices are typically easy to use, with no fiddly buttons or complicated on-screen instructions. Once set-up, users can retrieve information from the Web, select entertainment, contact friends, and make restaurant bookings without lifting a finger. A retirement facility in San Diego is reportedly experimenting with the devices as a way to support assisted living. Residents use voice assistants to set alarms, call their families and check the weather forecast. One resident, a 79-year old with hand tremors, finds it easier and faster than screen-based browsers to use his device to search the Web.⁶ Similar UK trials are taking place including in Hampshire where the technology is used to support adult social services and help address social isolation.⁷

Families have found the devices provide a way to reduce screen time. Random questions are answered and particular music played without everyone reaching for their smartphones. Google has suggested its smart speakers can help address dependency on smartphones by providing accessible alternatives, reducing the possibility of digital distraction and the impact this can have on sleep.⁸ According to a US survey by Edison Research, over two-fifths of users purchased a voice assistant in an attempt to cut down on screen time.⁹

Box 1: Scepticism

Some are sceptical about the disruptive potential of smart speakers. Technology consultant, Rob Blackie, points out that they are mainly used to select music, set alarms and answer questions – “hardly suggesting a fundamental change in people's behaviour”.¹⁰

Even if predictions about the speed of growth are overblown and their disruptive effect is exaggerated, the devices represent a new trajectory of technological development. Simply using them for web searches is a significant change to the way many people access information online. They also present technology companies with a new way to gather data which can drive further innovation, but may seem intrusive. This is a developing technology and we cannot be certain about the impact it will have. As a Microsoft report states: “This is just the beginning”.¹¹

6 Woyke, E. (9 June 2017) *The Octogenarians who love Amazon's Alexa*, MIT Technology Review.

7 See www.local.gov.uk/hampshire-county-council-pushing-boundaries-using-amazon-echo

8 See *Reduce Digital Distraction*, <https://wellbeing.google>

9 *The Smart Audio Report (Spring 2018)*, NPR and Edison Research.

10 Blackie, R. (20 February 2019), *The case for smart speakers being overhyped*, thedrum.com

11 Olson C. and Kemery K. (2019), *Voice Report*, Microsoft.

Future outlook

The capability of voice assistants is sure to improve. Users can expect more accurate word detection, more natural interactions, and wider 'skillsets' as both inbuilt functionality and app support is enhanced.

A little over a third of users already have the assistants set-up to interact with other smart devices. This is likely to change as more connected devices enter the market. An increasing range of hardware comes with voice assistants already integrated, including smart TVs and kitchen devices.

Voice assistants are also set to be used in a broader variety of contexts, including in the commercial sphere and in the provision of public services. Amazon is actively marketing products for pre-installation in hotel rooms and rental accommodation in the US. This, however, raises new concerns, particularly because of the users of the device being different from its owner, and the person using the space in question not necessarily having control over whether the device is enabled.

Services, like banking, delivered through voice assistants via smartphone apps or traditional phone calls may also become more common. Benefits of a more spiritual nature have also emerged. The Church of England launched an Alexa Skill in 2018 which can read a prayer for the day, say Grace or provide details of nearby churches.¹²

¹² See: www.churchofengland.org/our-faith/try-church-englands-alexa-skill

3. Data collection, use, and privacy

Despite their popularity, voice assistants have been plagued by news stories exposing the more unnerving features of this new technology. This has included the ‘creepy laugh’ controversy – where devices were ‘mishearing’ the command to laugh and seemingly chuckling spontaneously.¹³ In a more serious case, but seemingly rare, a device apparently mistook a conversation as a series of commands and sent a recording of it to a user’s manager.¹⁴ Such concerns generally split into questions around how data is recorded, and what is done with that data once it has been collected.

Always listening?

Much public unease has centred around how smart speakers collect and use data – in particular, the impression that devices are ‘always listening’, and essentially sending a recording of everything that takes place in a household to the cloud.¹⁵ This does not, however, reflect the functionality of major voice assistants. Audio is typically recorded, stored, and analysed locally on the device until the wake-word is detected.

Media reports have highlighted the tendency of voice-enabled devices to wrongly detect the wake-word and begin recording. In such instances, data is erroneously collected and sent to the cloud for processing and storage (known as ‘false accepts’). This has led to users discovering unintentionally recorded snippets of conversations found stored on their accounts.¹⁶ While devices typically provide indications that they have started streaming audio (e.g. through flashing lights), the surprise expressed in the media would suggest that these are not always effective in notifying users. Furthermore, as voice assistants become embedded in other hardware it may become less obvious when such devices have been activated.

The true impact of false accepts on user privacy is dependent on the subsequent processing by service providers and the level of user control over the recordings. With the technology rapidly improving, the occurrence of false accepts is likely to decrease and less data is expected to be wrongly recorded.

13 Liao, S. (7 March 2018), *Amazon has a fix for Alexa’s creepy laughs*, theverge.com

14 Wolfson, S. (24 May 2018), *Amazon Alexa recorded private conversation and sent it to random contact*, The Guardian.

15 See, for example, www.cnn.com/2018/05/13/are-siri-alexa-and-google-assistant-spying-on-me.html

16 Fowler, G. (6 May 2019), *Alexa has been eavesdropping on you this whole time*, The Washington Post.

Training assistants with user data

Their ease of use makes smart speakers particularly popular among those who are less tech-savvy. Providing people, who feel left behind by technological developments, with a new way to interact with technology brings many benefits. But such people may also be less aware of how the devices work. Indeed, even those who are comfortable with new technology may not have grasped what sort of data is collected and how it is processed. The media reaction to revelations that recordings are stored and even transcribed by humans is likely to be reflective of a wider societal naivety.

This has not been helped by the apparent reluctance of technology companies to explain what happens to the data collected by the devices. Recent media stories about data being collected have encouraged the companies to be more open. For example, in April 2019, Amazon admitted that a sample of recordings are reviewed by humans as part of its work to improve speech recognition and natural language understanding systems.¹⁷ Similarly, in July 2019, Google confirmed that recordings collected by its speakers could be listened to by contractors after a batch of its Dutch recordings leaked. Reports suggested that some of the leaked recordings contained identifiable and sensitive information which could be linked back to the user.¹⁸ Workers listening to the recordings also claimed they lacked guidance on how to respond when they heard something sensitive or distressing.¹⁹

In an apparent response to the leaks and reaction in the media, Amazon has given users the option to disable human review of their voice recordings, and committed to greater clarity about its use of this software training process in future. Google and Apple have suspended the practice altogether in Europe.²⁰

Users are expected to be active participants in the development of this new technology. Analysis of their interactions (albeit anonymised and only snippets of conversations) will help companies to develop better products – even when the data recorded is due to a false accept. The recent reaction in the media to revelations about humans listening to recordings may serve to encourage developers to make this clearer to consumers who are relinquishing aspects of their privacy. Despite requirements under GDPR for privacy notices to be concise and transparent, they still tend to run to several pages and are therefore not particularly digestible for users.

17 Cuthbertson, A. (11 April 2019), *Amazon admits employees listen to Alexa conversations*, The Independent.

18 Paul, K. (11 July 2019), *Google workers can listen to what people say to its AI home devices*, The Guardian.

19 Day, M. (10 April 2019), *Amazon Workers Are Listening to What You Tell Alexa*, Bloomberg.

20 Hern, A. (5 August 2019), *Alexa users can now disable human review of voice recordings*, The Guardian.

Controlling what is shared

Some developers have sought to address concerns about data collection by giving users greater control over what is shared. Google's My Activity hub, for example, allows users to review their voice and audio interactions, delete them, and turn off the automatic saving of interactions altogether. Amazon has introduced similar functionality and enables users to opt out of sharing voice data for the future development of new features.²¹

However, users need to make a choice. Deleting recordings or switching-off the saving function may disable quality-of-life features such as the device learning how the user speaks and what their voice sounds like.

What happens with the data collected?

Despite more clarity around the collection of data via smart speakers, the use of that data for profiling or other purposes not connected to improving the technology is opaque.

Although voice assistants are not used currently to advertise, the data collected from them can provide insights into users that informs advertising on other devices or formats. Data collected through people's online interactions, from the content they view to the comments they post or like on social media, has been a key tool in the development of profiling and targeting techniques deployed by advertisers and online platforms. Activity online (e.g. web searches, music choices etc.) via voice assistants will add to the complex adtech ecosystem. In many cases, smart speakers also have multiple users. The devices therefore potentially enable those collecting the data to gather new levels of insight into people's households – from what music they stream, to the time they set their morning alarms.

The extent to which deleting interaction logs or audio recordings also deletes any derived data or profiling is also unclear and may vary between different platforms.

The potential for analysis of sentiment, health and personality

By enabling platforms to collect audio recordings of people's interactions, voice assistants could take data analysis and inferences made about users to new heights.

Sentiment analysis is the practice of inferring additional knowledge from speech beyond the content of the words themselves, such as tone or mood. Given the type of data collected by voice assistants and the

21 Goode, L. (29 May 2019), *Amazon debuts a new Echo show amid Alexa privacy concerns*, wired.com

predicted increase in users, the potential for large-scale deployment is enormous.²²

Analysis may reveal insights into people's health, levels of fatigue, or other aspects of their personalities, paving the way for new levels of targeting. Last year Amazon patented a new version of its Alexa which could detect when users are ill by analysing their speech and offer to sell medicine to them.²³ Filing a patent does not mean the technology is being developed, but it may reflect a direction of thinking.

The potential benefits of such technology range beyond simply enabling new types of personalisation and are no longer theoretical. Doctors in the United States are using a smartphone App which analyses users' voices and detects early signs of imminent depressive episodes.²⁴ Researchers are also interested in other medical applications including detecting changes in breathing rates.²⁵ But targeting based on such insight may feel particularly intrusive or even creepy for users. It would require caution and a good level of understanding from consumers.

As with screen based interactions, these considerations are part of broader discussions about the extent to which large platforms are aggregating ever greater volumes of data about individuals through their different online interactions. Some of these challenges are not entirely new. Mental health, for example, can be estimated via other data sources such as email messages. Concerns relate to the adequacy of accountability as well as the impact on competition. Indeed, CDEI is conducting a major Review into the use of online targeting technologies. It is considering how targeting works, the benefits and harms it poses, and exploring public acceptance of these.²⁶

The potential of privacy-first architectures

Tech companies that produce solutions designed to protect privacy may be able to strengthen their position in the market. There is an opportunity for developers to adopt privacy first architectures, prioritising users' privacy over data collection.

Advances in 'edge computing' are likely to allow data to be processed locally on the devices, rather than on the cloud. Google recently announced that it had shrunk its speech recognition and language understanding models to a size that could fit on a smart speaker device.²⁷ In principle, this could

22 We understand that no commercially available platform currently employs sentiment analysis of voice recordings.

23 Cook, J. (October 2018), *Amazon patents new Alexa feature that knows when you're ill and offers you medicine*, The Telegraph.

24 Knowles, T. (15 June 2019), *Voice monitor app identifies the early indicators of depression*, TheTimes.

25 Blakely, R. (19 June 2019), *Alexa will get help if it thinks you're dying*, The Times.

26 More information about CDEI's Review at www.gov.uk/cdei

27 See blog by Manuel Bronstein Vice President of Product, Google Assistant (7 May 2019) <https://blog.google/products/assistant/next-generation-google-assistant-io>

mean less data leaving a home for processing. Whether or not this reduces the amount of data developers choose to stream to the cloud remains to be seen.

Dr Hamed Haddadi at Imperial College London is researching ways to limit excessive data sharing and communication between connected devices and unknown servers on the cloud. He asks, “Why does the server need to know exactly when I put the kettle on?”. His team is working on developing technology so that instructions to connected devices are kept locally and not shared with the server. Separately, Dr Haddadi and his colleagues are also developing noise shields designed to remove the emotion from voice content – so that less data is shared and only the instructions to the devices are collected.

Use of data by third parties

Digital assistants typically support the installation of third-party apps or ‘skills’ that extend their capability, for example by providing news, advice or smart home control.

Most platforms do not supply audio directly to app developers, but do pass on the data necessary for the relevant interaction with the app. This can include personally identifiable data, although platforms say they only do so with the user’s consent. In some instances, platforms also share audio transcriptions with third party developers to help them understand the nature of the interactions with their app.

Installation of apps is designed to be handled automatically or with the minimum ‘friction’ possible on many platforms. It is not clear that users are typically aware of what data is shared with third parties, who these third parties are, and how they can control the data shared beyond the voice assistant platform.

The smartphone app ecosystem has already experienced instances of apps covertly sharing user data with business partners.²⁸ There is a risk that similar behaviour may take place via smart speakers. As developers build their app stores they are well positioned to consider ways in which to better communicate with users around what data is used and shared.

28 Statt, N. (22 Feb 2019), *App makers are sharing sensitive personal information with Facebook but not telling users*, The Verge.

4. The disruptive effect

Voice assistants and smart speakers may transform the way people interact with the online world and can even fundamentally alter users' relationships with technology. The personalities of the devices affect how users engage with them and build them into their lives. This may revolutionise the way content is consumed and potentially leave technology platforms in even more powerful positions. While the technology continues to develop, concerns continue to be more speculative and therefore more difficult to address.

A New Relationship with Technology?

While voice assistants may support people's interaction with technology, the anthropomorphising of them has triggered a range of concerns. Some people's relationship with technology, particularly those who are vulnerable, particularly children,²⁹ may change considerably. They may come to view the devices as almost sentient beings with whom they can form relationships. The risk is that such relationships could be exploited resulting in people passively accepting the responses and instructions they receive without considering their own safety and security.

As the technology advances, it is expected to become possible to have more engaging conversations with the devices. Voice assistants cannot currently participate intelligently in conversations about politics or personal health. They generally only respond to instructions and relate what they hear to particular contexts. Developers are working on supporting more fluid conversations, so that the devices can interact and create "something that could potentially understand anything", according to Dr Layla El Asri at Microsoft Research.³⁰

Studies have suggested that people are more likely to open-up with non-human devices and share intimate details about their lives. The challenge lies in creating dialogue systems with safeguards which ensure users do not share information which could be detrimental to them. A situation where a device is, for example, giving mental health advice to a user who is high-risk. If developed responsibly

29 The ICO published its draft *Age Appropriate Design Code of Practice* in May 2019. It states "If you provide a connected device then you need to pay attention to the potential for it to be used by multiple users of different ages. This is particularly the case for devices such as home hub interactive speaker devices which are likely to be used by multiple household members, including children..." (pg 74).

30 See Microsoft Research podcast, *Talking with machines with Dr. Layla El Asri* (20 Feb 2019) www.microsoft.com/en-us/research/blog/talking-with-machines-with-dr-layla-el-asri/?ocid=msr_podcast_lasri_car

however, it could provide considerable benefits, enabling people to access support when they need it. Yet there needs to be confidence that devices will respond in an ethical and responsible way.

Public services and smart speakers

As noted above, voice assistants present opportunities to support the delivery of healthcare. In July 2019 Amazon and the UK Government announced Alexa devices will use the NHS website (nhs.uk) to provide information to users seeking health advice.³¹ This will help to give users confidence that the information they receive is from a reliable source. Indeed, in this case the ability of the technology to select such information as a default option is beneficial.

A number of commentators criticised the NHS for entering into a collaboration with Amazon. The nature of the criticism tended to be twofold. Some raised concerns about privacy and the sharing of personal information such as those described in this paper. Others were concerned that the NHS was encouraging the use of a particular smart speaker over others, thereby increasing its market power.

The agreement between the two organisations consisted in the NHS enabling Amazon to use the conditions data from nhs.uk, which is currently available via an API to app and web developers, and edit it slightly to turn it into a format appropriate for spoken dialogue (eg bullet points were removed). It is a non-exclusive, and non-commercial arrangement so the content API can continue to be used by developers for use by other smart speakers. No patient data is involved. If other developers follow suit, more users can be confident they are receiving reliable information when making health queries.

Such a move by Amazon, and possibly others in the future, may encourage users to share sensitive information about themselves through their devices. This strengthens the case for clarity around what is done with the data collected – and how users can control what is shared. However, it would be a disproportionate response for public services not to engage with this new interface. It is sensible for public services – which already make information available online – to support efforts to enable it to be accessed through smart speakers.

³¹ *Amazon Alexa offering NHS health advice* (10 July 2019), BBC News Online.

Box 2: Are voice assistants sexist?

The majority of voice assistants (British English Siri is a notable exception) use a female voice by default and two of the most popular (Alexa and Cortana) have female names. There are concerns the devices are entrenching gender stereotypes, presenting females as servile and docile, comfortable in the role of personal assistants. With companies developing products designed for children it is important to be cautious about inadvertently perpetuating gender bias in young minds.

According to the authors of a 2019 report for UNESCO, the devices are passive in their responses – “assertiveness, defensiveness and anger have been programmed out of the emotional repertoire of female virtual agents, while...sympathy, kindness and playfulness remain – as does stupidity...”.³² There is research suggesting users may prefer assistants to have a female voice.³³ However, Apple’s decision to assign Siri a male voice by default in Arabic, British English, Dutch and French, suggests that doing so is not commercially damaging.

In an interview with the CDEI, an experienced technology developer does not wholly accept the claims of sexual bias and suggests more nuance is needed. He argues the devices, which complete complex mathematical problems and provide authoritative answers to difficult questions, are possibly challenging sexual stereotypes. If the devices were seen as male, he suggests different criticisms would be made, leaving developers in a no-win situation. Even asking the user to select the gender of the voice might not fully solve the problem of perpetuating gender bias. One option could be to assign devices with gender neutral voices, or ones that deliberately do not sound human. It is not clear that this would be practical or commercially viable.

However, while women continue to be under-represented in the technology industry,³⁴ discussions about the design and functionality of devices are unlikely to be sufficiently inclusive. Until such under-representation is addressed, charges of sexism will be difficult to refute. The industry has also appeared reluctant to respond to the recommendations set out in the 2019 UNESCO report which includes the suggestion that developers run audits to identify gender bias in voice assistants and set out strategies to address it.

32 Unesco (2019), *I’d blush if I could: closing gender divides in digital skills through education*.

33 See, for example: <http://macdorman.com/kfm/writings/pubs/Mitchell2010DoesSocialDesirabilityBiasFavorHumans.pdf>

34 See, for example: www.pwc.co.uk/women-in-technology/women-in-tech-report.pdf

Impact on choice

The current market for voice assistants is dominated by a small number of large technology companies. Some fear such domination is a step towards the closing of the open-web. If the voice assistants become the key way by which users access content online, the major technology companies, through their assistants, strengthen their position as mediators. Critics do not view voice assistants as devices to enhance choice, instead they provide a single gateway to online content. This is particularly the case as people are also likely primarily to use one voice assistant, rather than several devices from different companies.

When people use their devices to search for content, the results are often driven by commercial algorithms. A user searching for an audio story may be nudged towards a particular provider if the manufacturers have agreements in place with them. While such algorithms also affect people's on-screen experiences, the impact is exacerbated through voice assistants, when it is harder for a user to review options.

The same issue arguably exists in the context of items being purchased via the devices. Users may struggle to compare different products available. If voice search becomes a prominent way by which people buy items, the impact of being among the top or default recommended products by the voice assistant may have a significant impact on competition between providers of those products. Of course, a user may also be unlikely to make a purchase without first viewing the item on screen.

Impact on news consumption

Half of users say that they use their device to access current affairs.³⁵ Yet, research has found that less than a quarter could remember the brand that produced their daily news update. How content is attributed when received through a voice assistant is not straightforward. Few people take the time to personalise or change their news settings. Upon set-up many devices now ask the user to select their preferred news source. This at least puts the user in greater control, enabling them to choose their preferred news source at a particular moment.

Attributing content, however, does not solve the whole problem. Device manufacturers are becoming curators or editors of content. Google has already created an aggregated news service – providing content from a range of publishers. While this offers attributed content, the way in which a news feed is curated, including how the stories are selected and the order in which they appear, puts the device developers and their platforms in influential positions. It is likely to accentuate concerns about filter bubbles in which people's news feeds align with their own outlooks. This is a recognised problem on-screen.

³⁵ Newman, N. (November 2018) *The Future of Voice and the Implications for News*, Digital News Project, Reuters Institute for the Study of Journalism, University of Oxford.

But it might be more difficult, or at least more time consuming, for users of smart speakers to access alternative sources.

The nature of voice assistants means they work best when providing concise information. This works well when a user requests local restaurant details or cinema listings. When an answer is drawn from news articles and the facts may not be clear, this is harder to communicate. Audio answers explaining that some 'facts' are disputed may come across as long-winded. This tension between efficiency and attribution is yet to be resolved. However as the technology evolves it may become possible for devices to highlight where outlets are reporting different figures and, if appropriate, point users to screen-based content for further information.

Opportunities for outlets to design and deliver content in different and engaging ways should not be overlooked. The Guardian recently worked with Google to set-up a Voice Lab to experiment with interactive audio. In a blog the research team concluded: "The prospect of creating seamless multimodal interactions across devices used to feel like the realm of science fiction. Through smart-assistant platforms, using both visual and aural methods as inputs and outputs is becoming an exciting reality".³⁶

A challenge to search engines?

The drive to provide people with information quickly and in a way that is easily digestible, will have disruptive consequences for traditional online searches.

Commentators often quote the prediction that by 2020, half of all internet searches will be made by voice and a third will be performed without a screen.³⁷ Voice assistants are forcing marketers to think differently about how they present content online. On-screen users rarely look beyond the third result returned by a search engine. When responses are delivered via audio it is likely that it will only be the first. The competition to have content at the top of someone's search results is also going to become more fierce. Christi Olson, Head of Evangelism at Microsoft, believes the impact on internet searching could be more profound. She argues that "website content in the era of voice search isn't about keywords, it's about semantic search and building the context related to answering a question".³⁸ Where this leaves consumers is not entirely clear. Retrieving information and purchasing products may become quicker and easier. But comparing products or reviewing content from different sources may be more difficult.

36 Pennycook, J. O'Donnell, D. Coleman, S. and Massey, S. (12 April 2019), *Speaking to machines brings opportunities and challenges*, The Guardian.

37 See, for example: www.wordstream.com/blog/ws/2018/04/10/voice-search-statistics-2018

38 Olson, C. (April 2016), *Just say it: the future of search is voice and personal digital assistants*, campaignlive.co.uk

Marketeers must also confront the ‘invisible shelf’ problem. Neither Amazon nor Alphabet currently allow direct advertising inside conversational voice apps. Yet given the expectation that voice-based internet searching is set to increase exponentially, there is a strong motivation for marketers to reach people through smart speakers. It may also be a tantalising new revenue stream for the major platforms.

Earlier this year it was reported that Pandora Media was selling commercial time in streams playing exclusively on Amazon Echo and Google Home smart speakers, possibly setting the way for advertisers to target people as they use home assistants.³⁹ If the major platforms relaxed their current stance, it could transform significantly targeted advertising.

39 Ives, N. (23 March 2019), *Pandora Pitches Ads Targeted to Amazon and Google Smart Speakers*, Wall Street Journal.

5. Ideas for further action

While new opportunities and benefits presented by smart speakers continue to emerge, the concerns surrounding them cannot be dismissed. In the short-term it is likely that technological improvements will reduce problems caused by 'false-accepts'. Despite the commercial incentives on platforms to maximise data collection, new products, designed to enhance people's privacy, may also become more widely available, particularly if there is a consumer demand.

However, fears raised about the longer term impact of smart speakers require different approaches. They could, in part, be addressed by developers taking steps to be more inclusive, responsive and open as they innovate. Recent and relevant Government activity in this area has focused on improving the cyber security of consumer Internet of Things devices, including smart speakers. This is part of efforts by the Government to encourage manufactures to make greater efforts to build cyber security within their products and help protect consumers' privacy, security and safety.⁴⁰ Furthermore, when it comes to concerns about attribution and assessing the reliability of content consumed, the Government, in its Online Harms White Paper, suggested platforms could be required to provide diverse news sources, and promote those which are more reliable in algorithms ahead of those that are less so.⁴¹ A similar approach to smart speaker providers could ensure alignment between this new technology and screen-based platforms.

Below we highlight further steps for consideration which are designed to address concerns highlighted in this paper:

- Industry and technology developers should work to explain in clear and accessible formats how the devices work, what data is collected, where it is stored, and how it is used, including any user profiling. They could adopt innovative approaches to doing this, including using the devices to succinctly explain privacy and data sharing policies and practices. The users' role in providing data that supports product development and innovation should be communicated clearly, including the possibility of human review of recordings, and any relevant privacy controls.

40 In 2019 DCMS ran a consultation as part of its Secure by Design work on their regulatory options for Internet of Things products. The aim of the regulation is to ensure products comply with global industry security standards. As part of this work the Government is considering initiatives to provide consumers with greater information on the cyber security of products.

41 *Online Harms White Paper*, DCMS, 2019.

- Users should be provided with more meaningful control over how data collected via smart speakers is used and shared, including with third party developers and contractors. New controls developed should be highlighted to users as they are introduced. The development of privacy-first architectures, including localised processing could be prioritised so that the amount of data shared to the Cloud can be reduced.
- Given the potential impact of voice assistants on privacy, the consumption of content and a user's ability to compare options, regulatory bodies (including ICO, OFCOM and the CMA) and the CDEI will need to keep abreast and gather evidence about the impact of smart speaker technology, as it evolves.
- Users should try asking their voice assistants questions about what data is collected and what is done with it. If they are not satisfied with the responses, they should consider adjusting their device settings or consider alternative devices.
- The CDEI's work on targeting and personalisation online will help to inform future thinking, particularly with regard to concerns about invasive profiling and how this might be addressed.

6. Conclusion

Smart speakers and voice assistants look set to have a disruptive effect on the way people interact with and control technology. They are already changing the way users consume content and are able to assess the reliability of the information they receive. This poses challenges as well as opportunities with such devices becoming key gateways to the online world.

An exploration of the concerns raised about smart speakers highlights the role consumers currently play in providing data which is used to support innovation. It is not clear that the public are aware of this and reaction in the media when examples of audio recordings are exposed suggests it comes as a surprise.

Equally, public acceptability around the potential use of audio data to profile and target users requires further exploration. The CDEI has a role to play in promoting such a conversation to consider these issues through its work in supporting ethical innovation. But most importantly, understanding what happens to data collected by voice assistants is not easy. Developers could be less opaque and consider ways to communicate better with consumers.

This might include adopting audio and visual formats to make traditional terms and conditions and privacy notices more engaging. Improved communication would enable consumers to make more informed choices and consider the trade-offs they are prepared to make, helping to address issues of trust. Recent scandals exposing the way data has been used and shared should serve as a wake-up call for platforms to act with more transparency and be proactive in explaining what data they are collecting and what is being done with it. There is a risk that society is moving towards a state where the devices are common place, but key questions around the data they collect and how it is used remain unanswered.

Smart speaker FAQs

What is a smart speaker?

Smart speakers are voice-enabled assistants – internet-connected software which responds to voice commands to provide content and services, interacting with users via digitally-generated voice responses. There are four major smart speakers available in the UK: Amazon’s Alexa, Microsoft’s Cortana, Google Assistant, and Apple’s Siri.

How many people use smart speakers?

The use of smart speakers has soared recently with at least one in 5 homes in the UK estimated to be using them.

Are they always listening?

Public unease has centred around the impression that devices are ‘always listening’. This does not reflect their functionality. At any given moment, the device holds a continuous ‘buffer’ of the last few seconds of sound recorded from its environment, which it scans for the wake-word. Only once the wake-word is detected does the device begin recording and streaming audio to the cloud for analysis and storage.

What are the benefits of smart speakers?

From setting timers while cooking, to calling friends without pressing any buttons, the devices enable multitasking and support people’s busy lifestyles. But for those who have a disability that makes using a keyboard or screen interface challenging, the technology may be life-changing. Devices could also be used to support the delivery of public services, including making information on the NHS website available to users seeking medical advice, as well as supporting residents in social service care facilities.

What are the concerns raised about smart speakers?

Public concerns focus on the seemingly intrusive aspects of the devices and the use of the data that has been collected. Questions have also been raised about their longer-term disruptive impact on the consumption of information and user profiling. Voice assistants, with human-like personalities, may change people's relationship with technology, particularly children and those who are vulnerable. Device owners may come to view them as almost sentient beings with whom they can form relationships. This raises the risk of exploitation, with people passively accepting the responses and instructions they receive without considering their own safety and security.

What is their impact on the online world?

Smart speakers present new and engaging ways for people to consume material online. However, a shift away from screen-based content may leave the major platforms in more powerful positions, with their devices becoming key gateways to the web. It remains to be seen whether this trend will negatively affect the quality of content that people consume, but regulators should pay close attention to it as an emerging issue.

What data do they collect?

The devices record audio data to process the user's request. It recently emerged that this can be reviewed by humans who are working to improve the technology. By enabling platforms to collect audio recordings of people's interactions, voice assistants could take data analysis and inferences made about users to new heights.

How is the data used?

The data is primarily used to process the user's request. The same data may be used for training purposes to support future product development. Given the type of data collected by voice assistants and the predicted increase in users, the potential for large-scale deployment of sentiment analysis is significant. This could bring benefits – for example for the delivery of health services – but needs to be done in a way that is transparent and trustworthy.

How will they affect choice and competition?

The current market for voice assistants is dominated by a small number of large technology companies. Some fear this is a step towards the closing of the open-web. If voice assistants become the key way by which users access content online, the major technology companies, through their assistants, would strengthen their position as mediators.

What can be done to address the concerns?

Over time, the likelihood of false-steps and data being inadvertently recorded is likely to decrease as the technology improves. However, developers could take steps to be more inclusive, responsive and open as they innovate. Regulators, including the ICO, OFCOM and CMA will also need to keep abreast and gather evidence about the impact of smart speaker technology as it evolves.

About the CDEI

The adoption of data-driven technology affects every aspect of our society and its use is creating opportunities as well as new ethical challenges. The Centre for Data Ethics and Innovation (CDEI) is an independent advisory body, led by a board of experts, set up and tasked by the UK Government to investigate and advise on how we maximise the benefits of these technologies.

The CDEI has a unique mandate to make recommendations to the Government on these issues, drawing on expertise and perspectives from across society, as well as to provide advice for regulators and industry, that supports responsible innovation and helps build a strong, trustworthy system of governance. The Government is required to consider and respond publicly to these recommendations.

We convene and build on the UK's vast expertise in governing complex technology, innovation-friendly regulation and our global strength in research and academia. We aim to give the public a voice in how new technologies are governed, promoting the trust that's crucial for the UK to enjoy the full benefits of data-driven technology.

The CDEI analyses and anticipates the opportunities and risks posed by data-driven technology and puts forward practical and evidence-based advice to address them. We do this by taking a broad view of the landscape while also completing policy reviews of particular topics.

More information about the CDEI can be found at www.gov.uk/cdei and you can follow us on twitter @CDEIUK

**Centre for
Data Ethics
and Innovation**

The Centre for Data Ethics and Innovation is an Expert Committee of the Department for Digital, Culture, Media and Sport. It is led by an independent board of experts and the views and statements of the CDEI do not represent government policy.

Design and typesetting by Soapbox: soapbox.co.uk