arm | NORTHSTAR

New Electronic Friends

Machine learning has found its

Everyday interactions with our devices can be frustrating. On keyboards, the @ symbol has a habit of moving location depending on where it was made and who designed it, no two TV remotes seem to be the same and there is a host of household electronics with non-intuitive controls. These differences, however small, make it hard to switch gears seamlessly between devices.

It's one of those truths you encounter every day and wish it would change...

Contents

Everyday interactions with our devices can be frustrating. On keyboards, the @ symbol has a habit of moving location depending on where it was made and who designed it, no two TV remotes seem to be the same and there is a host of household electronics with non-intuitive controls. These differences, however small, make it hard to switch gears seamlessly between devices. It's one of those truths you encounter every day and wish it would change...

Jem Davies
Fellow and General Manager of Arm's Machine Learning Group

This is a problem the industry has recognized for some time but the advent of machine learning (ML) technology offers a way forward. ML will ultimately allow us to interact with devices in a more consistent and human-centric way. The most obvious answer is to allow devices to understand human voice and gestures. But to be successful, this needs to happen with a near perfect degree of accuracy. So, while we have seen early voice and gesture capabilities on devices, most lack the accuracy and naturalness to be fully effective. ML will change that by allowing machines to recognize words, objects, gestures and motion as well as any human can, or possibly even better. It means the keyboard, dial or switch may not be as essential as it has been.

We already see glimpses of the future ML technologies can deliver as our smart TV suggests a show or our smartphone predicts what we're going to type, or automatically classifies images into categories for us.

This new survey by Northstar helps us to understand how people are reacting to these latest capabilities. It focuses on voice assistants, as one of the most visible examples of ML, and asks what users might want from new products in future? Also, which other devices could be powered by ML and what might it enable them to do? It is clear from what we've seen that new handsfree products have already gained a foothold, and expectations are starting to push ahead of what is possible today.

For Arm, enabling ML at scale in any class of device is a great challenge and in February 2018 we introduced Project Trillium, the next step in our broad enablement of ML. It represents a new wave of advanced ML and object detection hardware and software technologies that will lead to advanced products with increased compute capabilities. That might be a dive mask able to tell you what you are looking at – identifying fish and spotting danger lurking in a reef – or a city-wide camera system able to identify a person in need of help or spotting overflowing recycling bins.

We see a growing awareness and excitement around what machine learning will enable in the future and its ability to positively impact the way we live. While there is some frustration and concern among consumers, this survey shows people are beginning to look past the current limitations of technology to more advanced devices, interacting with us in the most natural of ways. I hope you find the insights in this report interesting.

2. Results from Global online survey

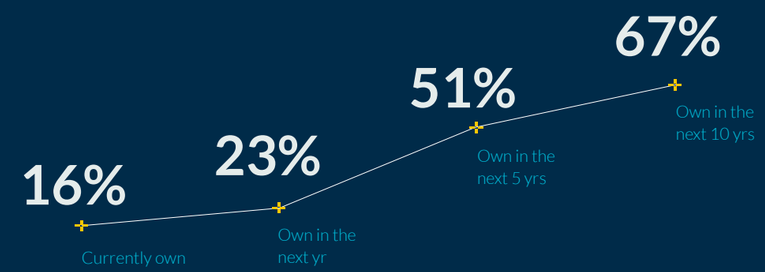
Advancing intelligence: understanding of machine learning today

- Phone knowing what you will type
- Voice assistants
- Netflix recommendations
- Amazon recommendations
- Car suggesting gas stations

Top 3 smart speaker uses

1. Entertainment, Music, Radio
2. News, Traffic & Navigation
3. Timers/ Alarms

The rapid rise of smart speakers will continue



Despite high satisfaction there are some frustrations

Voice assistants can be positively transformational

Benefits

- Less chance of human error
- Being able to identify illness and ailments
- Increase in productivity

Drawbacks

- Humans become reliant on machines
- Taking away jobs from humans
- Data privacy and security

3. The Wider Context of Voice Assistants

Things are getting smarter

Voice assistants are a high-profile example of how the exponential growth of artificial intelligence (AI) is re-defining human-machine relationships. Yet, they are only a single element in a much broader wave of transformation that will ultimately affect all layers of society and business.

Kevin Kelly, in his book The Inevitable, argues "it is hard to imagine anything that would change everything as much as cheap, powerful ubiquitous artificial intelligence". He predicts that AI and machine learning (ML) will embed intelligence in almost all devices and machines.

There is no doubt today's devices are far smarter than they were even a few years ago, and the momentum for intelligent technology is increasing. Globally 60% of people believe that in the next five years, personal devices will be far more intelligent than they are now.[1]

Manufacturers are providing, and users want, ever smarter devices. People want devices to learn their habits and better anticipate their needs. This is particularly true for smart home appliances and speakers, but it also applies to smartphones, tablets, cars and laptops.

This convergence between manufacturers' ambitions and customers' expectations is a clear signal that the rapid growth of ever-smarter devices will continue.

Figure 1: *How much more intelligent do you think your personal devices will be in five years? (% select)*

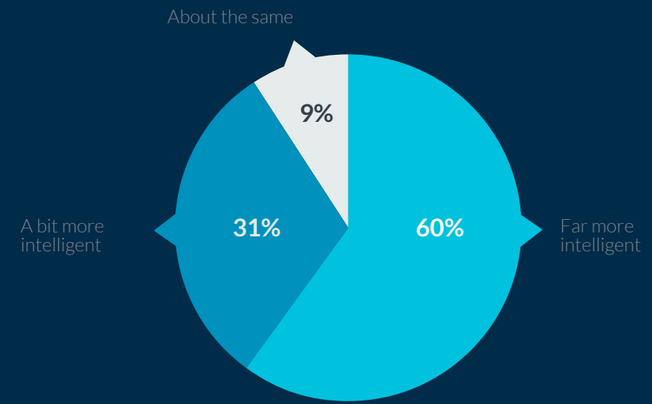
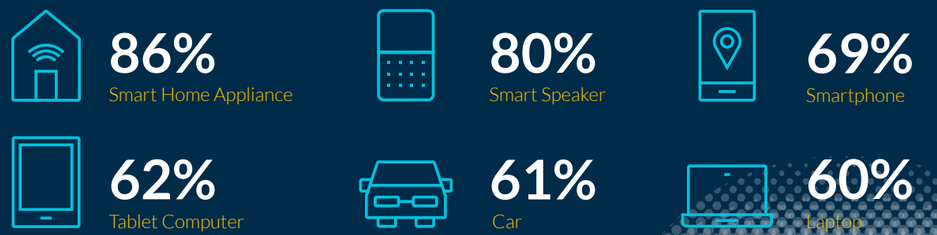


Figure 2: *Do you want the following devices to learn your individual habits so they can better anticipate your needs? (% select "Yes")*



4. Voice Assistant Ownership and Uptake

The year of the talk machine

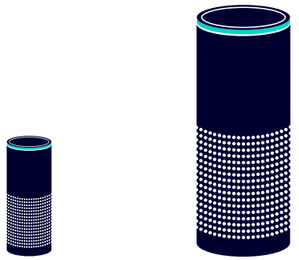
Perhaps the most visible example of ML is the voice assistant – a device that employs keyword spotting to trigger voice-activated actions, entertainment and information searches.

Voice assistants entered the mainstream in 2017 and smart speakers are now the fastest growing consumer tech product in history.

During the first nine months of 2017, 17.1m voice assistants were shipped worldwide. The December holidays and festive celebrations added an additional 16.1m devices in the last quarter of the year, with Amazon Alexa and Google Home topping the App Store charts on Christmas day.[2]

Our survey showed around one in six people already own a smart speaker. But this should rise significantly over the coming years, with two thirds of non-owners saying they intended to buy one over the next ten years.

Figure 3:
Which of the following do you currently own? Do you see yourself owning a Voice Assistant in the future?



16%
Ownership Now

67%
Ownership In The Next Ten Years

5. How voice assistants are used today

Habitual change

The table below shows how smart speakers are currently being used. Typical tasks include playing music, asking for news headlines or setting alarms and timers.

It is estimated that speech recognition is three times faster than typing, so there is a clear speed benefit to using voice assistance technology.[3] But more profoundly, speech is the natural form of human communication and is easier, and often more accurate, than typing.

Now technology has caught up, should we expect users to increasingly choose voice over fingers?

Asked whether they would prefer to fully control devices such as smartphones and laptops by voice control, around half of users globally would prefer this method over physical interaction. This suggests a large appetite for voice assistance outside of the home and away from smart speakers. A nod to a future where the predominant form of human-machine interaction is voice.

Figure 4: *How do you use your voice assistant?*

56%
Entertainment, E.g. Music, Audio Books, Launching TV Apps

35%
Web Browsing

54%
News, Traffic or Navigation

31%
Creating Shopping and To-Do Lists

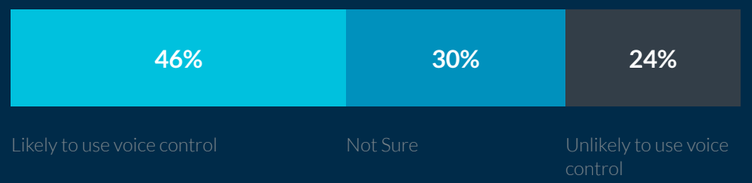
46%
Setting Timers and Alarms

30%
Home Automation (Lighting, Appliances etc.)

38%
Voice Calling/ Messaging

21%
Companionship/ General Conversation With The AI Assistant

Figure 5: *If you could fully control devices such as smartphones and laptops via voice control, how likely would you be to use that option more than physical interaction?*



Evolving expectations

With such a rapid rise in popularity, we are not surprised to see that three quarters of smart speaker owners are satisfied with their new talk machine.

We are currently seeing high levels of satisfaction contrasting with what appears to be rather basic usage behaviour – playing songs, listening to the news headlines etc.

As with many new and exciting technologies, the smart speaker appears to be experiencing a grace period where people are impressed by even the most basic of functionality (the 'Oh wow, I am talking to a machine' effect). A symptom of which is the wealth of strange questions that Alexa gets asked daily, as customers try-out their new novel technology.

It will be interesting to see how user expectations rise over the coming years as the technology becomes more widely adopted and the novelty-factor wears off.

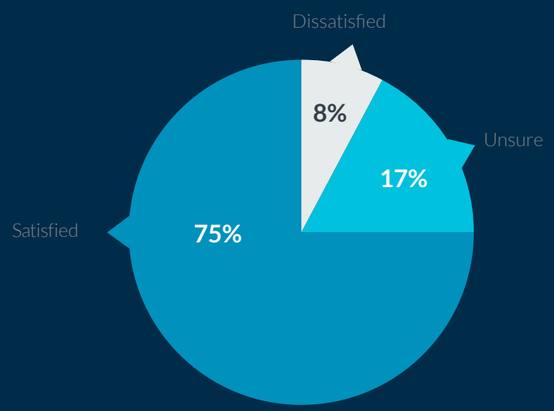


Figure 6: *When using your voice assistant, how satisfied are you with your device's ability to react appropriately to your requests?*

Teething issues

While overall satisfaction is high, there is still a significant minority of owners who are dissatisfied with their device's ability to react correctly to their requests. A quarter of users also say their devices often don't always understand them, highlighting the need for higher performance.

The biggest cause for concern however is that users don't want their smart speaker to rely on an internet connection to work. Other responses suggest this is about reliability, but also users not wanting to constantly share data externally.

The nature of ML algorithms means the accuracy of voice assistants will keep improving as they learn. The error rate of Google's speech recognition for example dropped from 8.5% to below 5% in less than a year.[4] That means Google now gets 19 out of 20 words right when it's listening to you talk.

While there are some early teething issues, advancement in performance and accuracy is rapid and so expect these concerns to become less of an issue moving forward.

Do you really know me?

There are two frustrations that, although not shared by most owners, point in the direction of an increasing set of user expectations. A quarter of people get frustrated that they cannot enter into a full conversation with their devices as they would with a human. Some 14% of responders indicated they didn't feel their device really knew them, perhaps indicating they would like that level of intimacy to change.

Before long, users will expect a seamless conversational experience, more than the largely transactional relationship that currently exists. And through the application of ML, talk machines will be expected to proactively anticipate the needs of their user, or in other words, get to know you better.

The implications of this hyper-personalized anticipatory design are significant. The rapid uptake of voice searches for example is changing the face of Search Engine Optimisation (SEO) as we know it, as businesses rush to ensure they feature at the top of the list when people search by voice.

And could this predictive behaviour even threaten our ability to make our own choices? Could it spell the end of chance and serendipity? Or conversely, will it just make our lives easier if devices could better predict what we might want to find or do?

Figure 7: *What frustrates you most about your voice assistant? (% select)*

34%
It has to be connected to the Internet

27%
It rarely understands me properly

24%
That I can't answer back to its reply

21%
It's not capable of doing the things I want it to do

19%
The delay in response

14%
I don't feel like it really knows me

6. How will voice assistants be used in the future

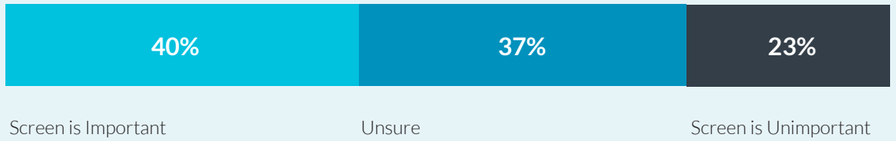
Next-gen voice assistants

Over the next year we will start to see the next generation of smart speakers, many of which may include screens, allowing new ways of interacting with these devices.

Google recently announced a series of deals with companies including JBL, Lenovo, LG and Sony, in which their Google Assistant technology will be embedded in speakers that all contain screens as well. This allows for seamless interaction with video and image platforms all controlled via voice, expanding the level of functionality significantly.

But there is divided opinion around the importance of including a screen, with almost two fifths of consumers unsure of the added benefits. It will be interesting to see how this opinion changes over the next year with the more widespread adoption of voice assistants with screens.

Figure 8: *How important do you think a screen is in interacting with a Voice Assistant? (% select)*

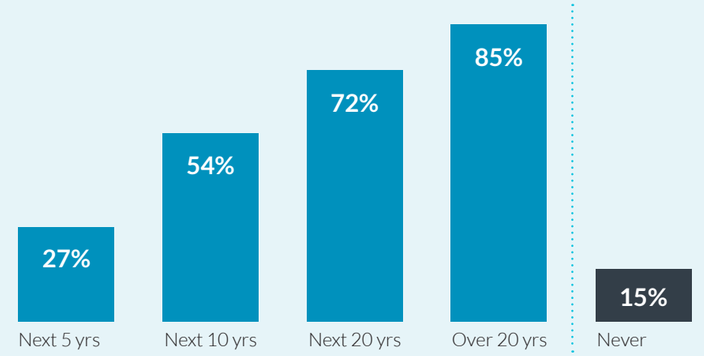


The race to sound human

One technological milestone that is still to be achieved is the development of an assistant whose voice and communication is indistinguishable from that of a human.

The global opinion is that this will be possible and more than half of respondents believe it will have been achieved in less than ten years. But even that might be slightly pessimistic. In 2017, DeepMind's WaveNet team developed a new convolutional neural network that improved performance significantly. It is now being used in the Google Assistant, making interaction that bit more human. [5]

Figure 9: *How far into the future do you think speaking to a voice assistant will be indistinguishable from speaking to a human?*



From assistant to companion

Although voice assistants are typically being used in a functional way, more than one in five users said they also use their device for companionship and general conversation. This poses an interesting question about just how evolved human-machine relationships might become. In the future, will people grow to think of their voice machines as friends, rather than assistants?

For lonely people, this companionship could be extremely beneficial. An elderly person living by themselves for example could benefit from a voice assistant in many ways. Not just by asking it to complete tasks or set reminders, but by engaging it in conversation to stave off feelings of isolation in the same way they might talk to a pet.

7. The overall impact of voice assistants

Cautious optimism

In a previous Northstar report sponsored by Arm on artificial intelligence, AI Today, AI Tomorrow, we saw a cautious optimism towards AI. And the same is true for voice assistants. Most people around the world are generally positive, though there is a significant minority with some reservations.

The primary worry is that voice assistants will contribute towards humans becoming too reliant on machines, which is a wider consideration when it comes to AI and ML.

Voice assistants also raise data privacy concerns, as most users would prefer to have sensitive data stored locally. It is interesting to note they are even willing to accept slightly lower performance from their device in return for greater security.

Figure 10: *Overall, how do you feel about the emergence of voice assistants?*

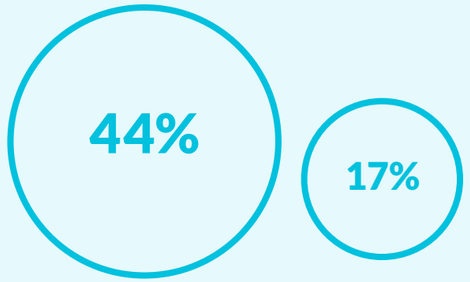


Figure 11: *What do you perceive to be the negatives around voice assistants becoming a main part of society?*

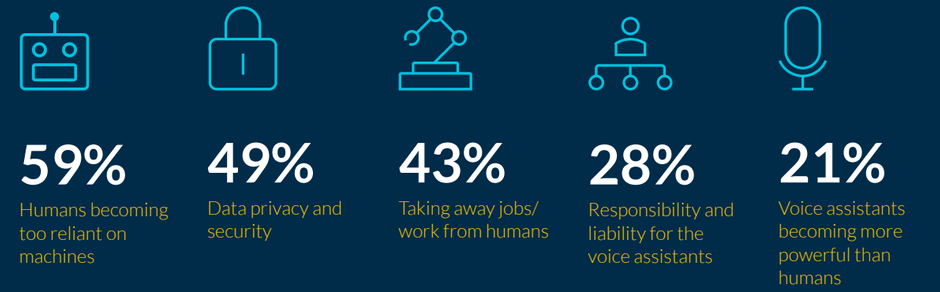
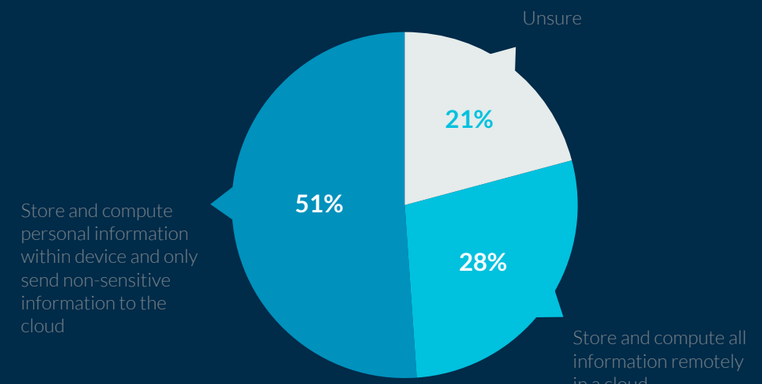


Figure 12: *Would you prefer the information your device learns about you to be stored locally on your device or uploaded to the cloud for processing?*



Moving to the edge

There are technologies coming now that will mean that there doesn't have to be a trade-off between performance and privacy. The main way is to introduce greater ML capability at the edge – processing information where it is gathered rather than in the cloud. It means many tasks will be dealt with by devices locally. When a cloud-connection is required, for an internet search for example, the task can also be carried out more efficiently and potentially more discretely.

Arm recently launched a suite of new ML technologies under Project Trillium. It is a plan to distribute ML capabilities across all device categories, from single sensor products through to data center infrastructure and all points in between.

This level of compute performance will bring other advantages such as helping to overcome bandwidth and power challenges, or issues of reliability and latency.

Focusing on latency: there will be many functions of smart devices where response delays from performing ML processing in the cloud will be unacceptable. This will certainly be the case if we are to expect real-time conversations with our devices.
The need for immediacy is heightened when looking at areas such as autonomous driving, where any delay in processing could be a safety issue.

A positive development

Despite some concerns, the overall outlook is one of optimism. Voice assistants are being perceived as potentially more than just helpers.

The top perceived benefit is limiting human error. An indication of the growing amount of trust in the technology.

And while smart speakers are being adopted in homes across the globe, voice assistance is also being used within the workplace. It may ultimately offer workers a chance to augment what they can do, and improve decision-making, by collaborating with a virtual colleague.

8. So what next for voice assistants?

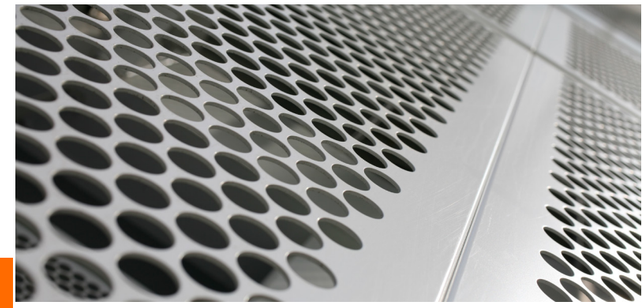
The rapid rise of smart speakers is indicative of a consumer appetite for smarter devices and voice assistance technology. According to the results of this survey, the momentum is set to continue over the coming years with smart speaker adoption rates continuing to soar.

Werner Vogels, Chief Technology Officer and Vice President of Amazon believes that by the end of 2018 "voice will have changed the way devices and applications are designed" and that it "will be on the way to becoming the primacy interface by which we engage with technology and the world around us". [6]

So beyond smart speakers, could the next few years prove an inflection point where voice assistance technology achieves ubiquity, embedded in a much broader spectrum of devices and machines? It seems so from looking at this year's Consumer Electronics Show in Las Vegas, where a host of devices powered by machine learning became star attractions. And the results of this survey suggest the consumer appetite exists.

So, from a rising number of single electronic voices, we are set to hear and see a global chorus of exciting new directions for advanced technology.

Hello future…



9. A demographic perspective

Age

Globally younger people are more willing for machine learning to be embedded in devices to make them smarter and more anticipatory of their needs. They are also more comfortable with devices storing and computing their personal information remotely in a cloud.

Ownership of smart speakers peaks among younger people, and is lowest among the over 50s. The youngest owners under 30 have higher expectations and as a result get more frustrated with the performance of their device. They are more likely to complain that their smart speaker rarely understands them properly, as well as being frustrated by response delays.

Region

Overall, we found that people in Asia tend to be the most positive and excited about the emergence of voice assistants, despite currently having the lowest ownership of smart speakers. Users in Asia are most open to the idea of their devices learning their behaviour and anticipating their needs and the most likely to look for companionship.

Although still mostly positive overall, Europeans are the most sceptical towards voice assistants. They are the most likely to not find any societal benefits and the most likely to think that voice assistants will never advance to a point where they'll be indistinguishable from speaking to a human.

Smart speaker ownership is highest in the US, with using devices for entertainment purposes particularly popular.

10. A note on the research

Methodology

This report is based upon insights unearthed from a global consumer survey. The survey was approximately 10 minutes in length and asked questions about Machine Learning, current Voice Assistant usage and applications for the future.

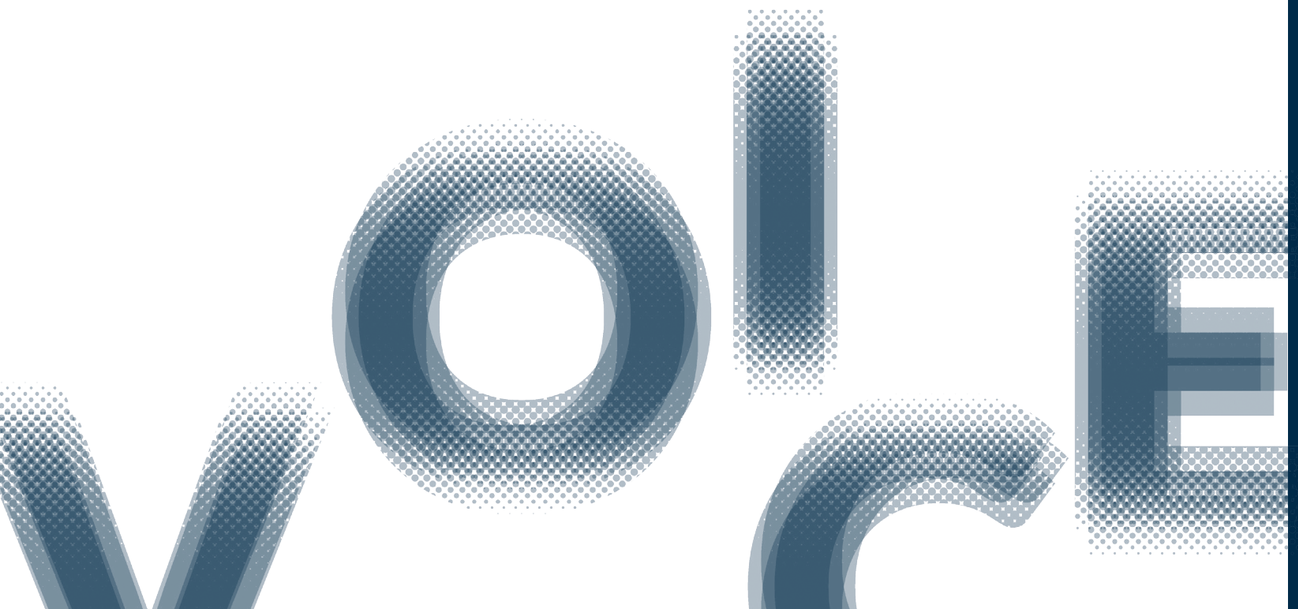
The research was conducted by Northstar Research London and sponsored by Arm.

How many people were spoken to?

2,133 consumers were interviewed from US, Asia Pacific and European regions through an online survey.

Who have we spoken to?

Representative sample in each market, with matching age profiles. There were no additional screening criteria.



11. References

[1] **Kelly, K, The Inevitable,**
New York: Penguin Books, 2016

[2] **Gibbs, S (2018),**
How smart speakers stole the show from smartphones
Available at:
https://www.theguardian.com/technology/2018/jan/06/how-smart-speakers-stole-the-show-from-smartphones
(Accessed: 20th February 2018)

[3] **Ruan, S; Wobbrock J.O.; Liou, K; Ng, Andrew; Landay, J (2016),**
Speech is 3x Faster than for English and Mandarin Text Entry on Mobile Devices
Available at:
http://hci.stanford.edu/research/speech/paper.html
(Accessed: 20th February 2018)

[4] **Google Developers (2017),**
Google I/O Keynote (Google I/O '17). [Online Video]
Available at:
https://www.youtube.com/watch?v=Y2VF8tmLFHw.
(Accessed: 20th February 2018)

[5] **Van den Oord, A; Walters, T; Strohman, T (2017)**
Available at:
https://deepmind.com/blog/wavenet-launches-google-assistant/
(Accessed: 20th February 2018)

[6] **Wogels, W (2018),**
Voice is about to fix our love-hate relationship with machines
Available at:
http://www.wired.co.uk/article/google-home-amazon-alexa-apple-homepod-voice-recognition
(Accessed: 20th February 2018)

arm | NORTHSTAR