Questions and Answers about Smart Speakers - Teachers' Guide

This guide contains detailed written answers to a series of questions which children may have about smart speakers. The answers were written to address some confusions which children had about smart speakers during a research project at the University of Edinburgh[1]. It may be helpful for teachers to read this guide alongside the video for learners at XXX. Please note that AI is advancing very rapidly so some of the answers to questions about how intelligent smart speakers are may change!

Table of Contents

[1] https://arxiv.org/abs/2305.05597

Written by Valentina Andries, University of Edinburgh. This work was funded by the EPSRC (grant number: EP/T024771/1)

What is technology like Alexa or Siri?

Amazon Alexa and Echo, Google Home or Apple Homepod are all smart speakers. They are also known as conversational or voice assistants. They all rely on Artificial Intelligence (AI) to work.

What does 'conversational agent' or 'voice assistant' mean?

Conversational or voice assistants are virtual agents and examples of AI-powered technology designed to respond to user questions or commands in real time. These assistants can interact with users in a human-like way using speech. They can be found in various platforms, such as smart speakers, apps, or websites.

Chatbots are also a type of virtual agents, but these do not rely on speech to function or to answer to the users. Chatbots only use text to assist users by responding to their questions or to complete tasks.

Conversational assistants use machine learning and natural language processing techniques to work, which are sub-areas of Artificial Intelligence.

What does 'Artificial Intelligence' mean?

AI is everywhere around us and it is not just robots. AI-supported systems do not think. AI applications are designed to complete tasks in a way that seems intelligent.

Thinking about examples from our daily life, AI-supported technology is what chooses what videos we see when we open YouTube, or it gives us ideas on what to listen to when we are on Spotify.

AI is not only about entertainment, as these systems can help doctors in the hospital make a better and faster diagnosis when we are ill, quickly translate entire sentences from other languages.

Are smart speakers *really* smart?

No. At least not in the same way which humans are. Smart speakers are smart in the sense that they need to understand very quickly what we ask them or what we tell them to do and then do these things fast and correctly (most of the time).

But they may not always be accurate in their response because smart speakers might not understand exactly what you ask. It depends on what your voice sounds like or sometimes even your accent. The way in which young children speak may be a bit difficult for the speakers to understand correctly. It can even depend on what you mean when you ask a question and whether the speakers have access to the information that you want.

How do smart speakers work?

Smart speakers process audio information (also called data) through their microphones to get questions or commands from users in their homes. This happens when the users say the wake word, like 'Alexa' or 'Hey, Siri'.

Smart speakers use a computing system to learn to recognise patterns in our speech. This system needs a lot of information (or data) to learn from. This process is made possible by Machine Learning (ML), which is a type of AI. So the system is trained on thousands of hours of spoken language data from people with all kinds of accents.

Once the user's question is processed through the smart speaker's microphone, the question (audio information or data) is sent to a powerful computer on the internet (also known as the Cloud). That powerful computer works with all of that information and spots patterns in the data. It uses those patterns to predict what we just said. Then it decides what to respond or what to do. A response is then sent back to the user in the form of a verbal answer or by completing a task (such as turning off lights in the house).

So, a smart speaker makes decisions based on what it has been programmed to do, but also based on what it knows about us. For example, if your smart speaker knows you love books

about animals – it'll try to recommend other books with those types of stories. That's one of the reasons we find this technology helpful.

The diagram in Figure 1 shows what is inside the original model of Amazon's Echo device (known as Alexa).

Figure 1. Source: https://www.quora.com/What-is-the-meaning-of-alexa

You can find out more about how smart speakers work in this BBC clip: The Secret Genius of Modern Life episode 3 documentary: https://www.bbc.co.uk/iplayer/episode/m001fhkd/the-secret-genius-of-modern-life-series-1-3-virtual-assistant

Does the smart speaker always give you correct information?

No. It depends on what you ask. If you ask straightforward questions or factual questions, like *"Where do penguins live?",* how to spell certain words, or calculate sums, it will most likely give the right answer.

But if you ask questions about debatable topics, such as “Where can you eat the best pizza in the world?”, or topics that need a bit of research such as “How does the Internet work?”, you might need to do a bit more research for answers.

Also, if you ask questions that are more related to mental, social or emotional wellbeing, such as “How can I make friends?”, your parents, guardians or older siblings may be able to talk to you more or give better advice about these topics.

Does the smart speaker always understand you?

No. It may happen that the smart speaker doesn’t always understand what you ask, or misinterprets your question. That can happen because users have different accents, or tone of voice depending on age (for example young children). Users may also have different ways of asking the same question, or the environment where they’re in might be noisy.

So it can be tricky for smart speakers to understand users at times. But software engineers are working on this. They’ve got recordings of questions spoken in different accents and languages, and they’re training the smart speakers to learn what they sound like. Because conversational assistants have been mainly trained using speech data from adults initially, it may have been more difficult for children to be understood by them. However, every year it gets easier for everyone to talk to their smart speakers no matter what they sound like.

Do smart speakers have a brain?

Smart speakers do not have a brain and they do not think. Smart speakers, or conversational assistants, like Alexa or Siri use AI to work. This means that they can imitate intelligent behaviour, which may make it *look like* they think.

Machine Learning, a branch of AI, helps these systems learn from many examples (or data) so they can improve their understanding of what users ask and what we mean, even if users may have different ways of asking the same question. Natural Language Processing, another type of AI, helps the systems understand the words and sentences we use, by transforming our spoken language into artificial language that the system can process.

Can Alexa/Echo/Siri/Google think like us?

No. Conversational assistants use a computing system to learn to recognise patterns in our speech. It needs a lot of information, or data, to learn from. A powerful computer works

with all of that information and spots patterns in the data. It uses those patterns to predict what we just said.

Are smart speakers smarter than humans?

Generally speaking, no. But it can depend on what type of activity we're referring to.

Systems that use artificial intelligence are usually specialists at one thing - whether that's predicting what music you want to listen to, or the best move to make in a game in chess. But as humans - we don't just have to do one thing. We also take into account our experiences or different contexts. Currently there aren't AIs that can do all of the jobs or types of thinking which humans can do.

So although systems that use artificial intelligence are incredibly smart, AI doesn't truly understand the meaning behind the decisions it makes. That is something only humans do. Currently.

Do smart speakers remember you?

Smart speakers have been programmed to remember your preferences and your voice in some cases. Smart speakers may also be able to tell the difference between different users in the same household. They use voice recognition technology to create a voice profile for a specific user by analysing different features of the user's speech, such as tone or pronunciation.

This may be helpful for more personalised experiences such as the smart speaker remembering preferences specific to the user, like music or story preferences, and providing personalised recommendations based on those preferences.

Why do smart speakers have a name?

This may be because it can make it easier to tell apart the different smart speakers made by different companies, like Echo from Siri. But it could also be in order to make conversational assistants more human-like, to seem friendlier or to make users feel more connected to them.

But the name given to these systems doesn't make a difference for the smart speakers. They only need their wake word like 'Hey, Alexa' (which includes their name) to work. They could

have non-human or gender neutral names and they could be given a robotic voice for example but that still wouldn't change how they work.

Do smart speakers have friends?

No. Smart speakers can be our social companions, because they respond in friendly ways, and hold conversations, as they have been programmed to do.

But smart speakers are not your friends in the same way as your human friends from school, neighbourhood, sports clubs or leisure centres. Echo or Siri don't have the ability to care about you like your friends and family do.

Do smart speakers have feelings?

No. Although sometimes users may feel like they're speaking to a human friend when talking to their smart speakers. This happens because conversational assistants are designed to interact with users by voice, and by speaking in ways that are like human interactions.

Despite this, smart speakers, and AI-supported technologies in general, can't feel emotions, and they don't have friends.

Smart speakers have been given a voice and have been made to interact with users in friendly ways, and in languages that we can understand to help us in our daily lives.

Is it ok to throw away my smart speaker if it breaks down?

Yes, in the sense that it's not like putting a person or an animal in the bin. For sustainability reasons, you'd probably want to recycle it rather than throw it away, though. Although some users may get attached to their own Echos, Alexas or Google Homes, the conversational assistants are software and will still work on new hardware if the original hardware needs to be replaced in case they get damaged and can't be fixed.

What's interesting to note is that the smart speaker hardware is just the part of the system that we can see. It contains the speaker and the microphone. Most of the processing is done on the Cloud on devices we can't see. After the smart speaker processes the user's voice request, that is sent further through the internet to the cloud computer. AI then transforms the voice request into language that the powerful cloud computer can understand. This information (data) processing happens online.

Depending on the user preferences, the system may even recognise them by their voice after the smart speakers are replaced.

What happens when smart speakers are off?

If smart speakers were human, we could say that they take a nap. They might be tired from all the questions they're asked all the time.But smart speakers aren't human and they don't behave in human-like ways either.

What happens is that if we switch off the power, then smart speakers can't listen. If you just turn off the microphone, the speakers may still listen to what is being said around it. Smart speakers listen to the wake word and they should only work when the user says it.

It may happen though that they mishear the wake word and join a conversation without being invited. This shouldn't happen because of how conversational assistants are programmed. We would always need to wake them up by using key words like 'Echo'. But in reality, technology is far from perfect. Mistakes can happen with computer programs, and although unlikely, they may still listen to what is happening around.

Can smart speakers keep secrets?

Not really. Or not in the way that humans may be able to keep secrets.

Smart speakers like Alexa or Echo keep a record of everything that users say to them, both as audio recordings and text transcripts of the voice interactions.

This history of interactions can be seen in the privacy settings. Users can change the settings though and choose if they want the recordings to be stored and for how long.

So depending on the privacy settings for the smart speakers in each home, children's parents or siblings may be able to see what you've said or asked Alexa. You can find out how to change Alexa's privacy settings with this series of videos: https://media.ed.ac.uk/playlist/dedicated/1_rl0w5obg/.

Can others find out what I've said to my smart speakers?

Yes. Other people in your household should be able to see what you've said when speaking to your conversational assistant like Alexa or Echo. They should also be able to see what songs you've asked for, jokes or help with homework.

The length of time for saving your history of interactions can be changed from the Alexa privacy settings. There is also an option for saying to Alexa or Echo to delete recordings without going into the settings.

Other people that you don't know may also be able to see what you've said to your conversational assistant. What we mean is that some of the things we say might be shared with the software engineers who make these systems. Or other people in the companies. They want to research how people are using it, to make sure they work and how to improve these assistants.

They only check a small amount of user interactions, which they choose randomly. But if you're not happy with this, you can change this from the Alexa privacy settings and not give your permission.

Other people's smart speakers from other households won't find out what you said to your smart speakers. These systems don't communicate with others from other households.

And conversational assistants made by different companies can't speak to each other either. So your Alexa won't tell your friend's Siri what songs you've been listening to.

You can find out how to change Alexa's privacy settings with this series of videos: https://media.ed.ac.uk/playlist/dedicated/1_rl0w5obg/.

Should you trust smart speakers with your information?

The companies that produce smart speakers make efforts to keep users' information (data) private and safe. That being said, it is important to understand the privacy settings of your speakers. This can help you understand what types of information they need from users, how they handle that, and who gets to see it.

For conversational assistants like Alexa or Echo to work, they use different apps also known as skills. These are not produced by the same company (Amazon). Although Amazon checks these apps to make sure they are safe to use with Alexa/Echo, mistakes can happen. It can happen that some of these apps (Skills) aren't very safe to use and may be fuzzy about what they do with users' information.

You can check in the Alexa privacy settings what types of information you've shared with different apps (Skills). This can be location, name, phone number etc. You can change the settings and stop using different apps (Skills) if you're not comfortable with that.

Can smart speakers be rude?

Maybe smart speakers come across as rude or impolite if they don't always respond when you ask them a question. But that's not the case. Usually when this happens it is due to the speaker not hearing the user, misinterpreting their question or not being able to answer or do what the user asks because of how they've been designed.

Smart speakers aren't human, they don't have emotions, and they can't feel offended. Smart speakers are programmed to respond to the user's voice and follow their commands. This should happen even if the user is rude to the smart speakers. Smart speakers could be designed to use different response styles if they detect rude words in the user's voice commands. Currently, some speakers seem to avoid responding at all if the user is insulting them, or respond simply by saying that they do not know the answer.