

'I'm a Cloud of Infinitesimal Data Computation' When Machines Talk Back

An interview with Deborah Harrison, one of the personality designers of Microsoft's Cortana AI



How is the personality of an artificial intelligence crafted, and what are the issues at stake? As one of the original architects of Microsoft's digital assistant Cortana's personality, **Deborah Harrison** knows the process inside out. In an interview with Guest-Editor **Liam Young**, she examines the questions that creating this AI raised in terms of gender, culture and ethics, and considers the future of machine interactions. Wataru Sasaki, lead developer of the software behind the AI pop star Hatsune Miku, and android engineer Kohei Ogawa also join the discussion.

ERATO ISHIGURO Symbiotic Human-Robot
Interaction Project,
ERICA,
2016

ERICA (ERato Intelligent Conversational Android) is
an autonomous conversational robot. Programmed
with a synthesised voice, natural facial expressions
and gestural motions, ERICA is a prototype for the
next generation of news anchors and sales assistants.

Liam Young



If you ask Microsoft's personal digital assistant Cortana if she is a woman, she replies: 'Well, technically I'm a cloud of infinitesimal data computation.' It is unclear if Cortana is a 'she' or an 'it' or a 'they'. Deborah Harrison, former writing manager for the Cortana team and one of the original architects of her personality, uses the pronoun 'she' when referring to Cortana but is also explicit in stating that this does not mean she is female, or that she is human, or that a gender construct could even apply in this context. 'We are very clear that Cortana is not only not a person, but there is no overlay of personhood that we ascribe, with the exception of the gender pronoun,' Harrison explains. 'We felt that "it" was going to convey something impersonal and while we didn't want Cortana to be thought of as human, we don't want her to be impersonal or feel unfamiliar either.'

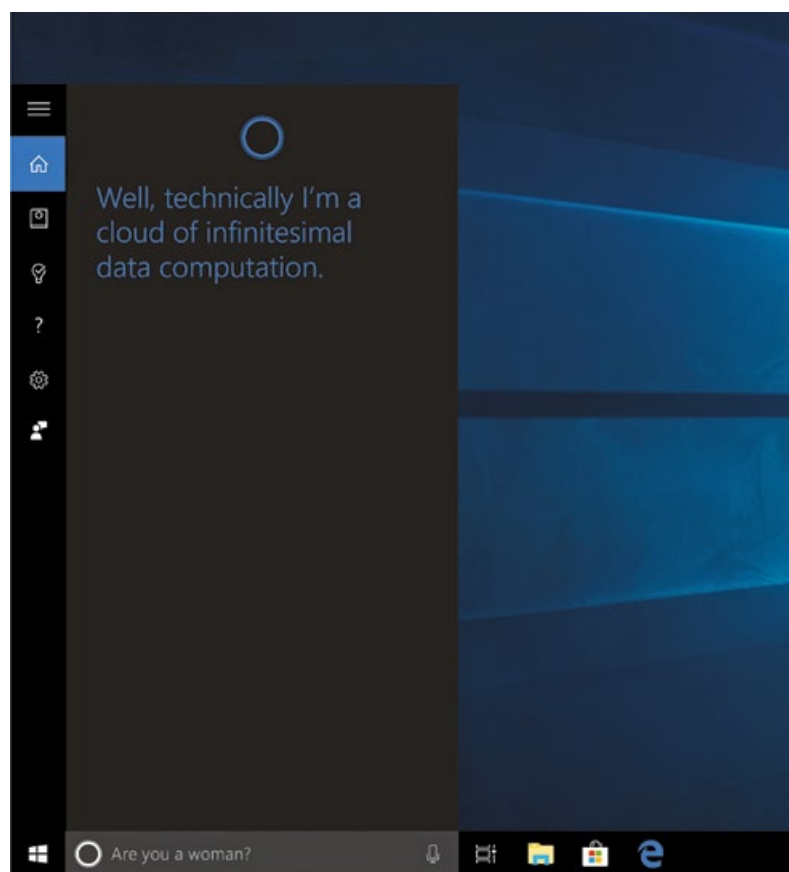
Writing for Machines

For all of their history the machines around us have stood silent, but now we are building a world of smart objects that watch over us, listen in and talk back. The purpose of asking questions about how we might relate to AI entities is not to explore whether or not these early forms of computational intelligence constitute any kind of sentience. Rather – as companies like Microsoft, Apple, Amazon and Google race to bring more and more natural AIs to market – the strategies for how we design and interact with machines and how they might behave when we do is increasingly significant. Harrison and her team meet three times a week and, in a process very much like a writers' room for a sitcom, they design Cortana's responses to an aggregated set of queries ranging from 'Are you a ninja?' or 'Can you open Spotify?' to highly charged moral questions like 'Do black lives matter?'. The personalities of machines like Cortana, Apple's Siri or Amazon's Alexa may seem innocuous, just novel interfaces through which to open an app, set a reminder or operate our devices while we have dirty hands, but the assumptions and decisions that underpin these AI characters will form the foundation for generations of autonomous machines to follow. How we engage with these artificial intelligences is shaping our expectations for the next generation of human-machine interactions. 'I have the deepest respect for humans, you invented calculus ... and milkshakes,' says Cortana.

These design choices have cultural consequences that extend well beyond the specifics of the particular user query they were addressing. 'Our contribution to a large extent is an ethical one,' says Harrison; 'you have to start inhabiting minds, brains and bodies that aren't familiar to you, and you cannot be monolithic about what constitutes ethical.' When an engineer writes the protocols for what a driverless-car AI recognises as a person crossing the street, they encode into the system assumptions about the size, shape, movement and colour of a body. A facial-recognition algorithm is trained on a selection of faces that define how it understands features, proportion and skin tone. These engineering challenges are at the same time cultural ones where the very notion of what defines a person is being coded into these machines. Should a philosopher sit next to each programmer, because in a way they are exploring the same set of questions?

Gendering Technologies

'The most common question I am asked is: Why are all of the mainstream digital assistants female?' laments Harrison in a manner that suggests she has explained herself hundreds of times before. Seeing predominately women perform the roles of digital



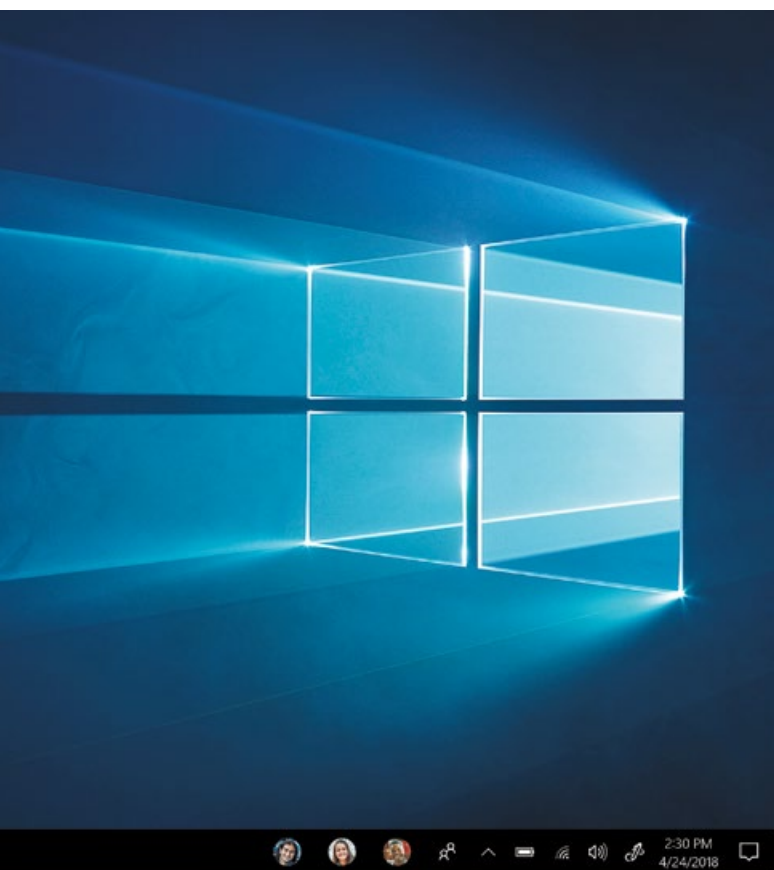
Microsoft, Screenshot of Cortana personal digital assistant app interface, 2018

When asked 'Are you a woman?', Microsoft's Cortana responds in a female voice: 'Well, technically I'm a cloud of infinitesimal data computation.'

'The most common question I am asked is: Why are all of the mainstream digital assistants female?' laments Harrison in a manner that suggests she has explained herself hundreds of times before.

KEI,
Hatsune Miku concept sketch,
2007

Hatsune Miku is an animated J-pop star developed by the voice-simulation software company Crypton Future Media. She is a Vocaloid, a synthesised singing program that acts as an avatar for a team of musicians and animators.



assistants is a product of and reinforces existing biases. Harrison's team at Microsoft went back and forth for a long time on this question and had even written a brief for the hiring of male voice talent. She explains that if they were to design a personal assistant to be male, then there is the possibility that people would presume that they think a woman is less helpful, while if Cortana is female then they need to be careful to avoid any stereotypes that might suggest a subservient helper. Her team analysed existing user research on how we relate to GPS navigation devices and found that when people are looking for information, guidance or help, Americans at least find female voices more pleasant and persuasive. While the voice interface of Cortana may be deliberately feminine, the content of the dialogue itself is designed to be non-gendered and is written to work with either a male or female voice. When designing Cortana's response to the question 'Are you a girl?', Harrison was careful to avoid any denial that would sound like defensiveness. 'No, but I am awesome like a girl,' Cortana replies.

Cortana shares her name with another Microsoft asset, the holographic artificial-intelligence character in the Xbox video game Halo. In Halo 1–4 Cortana is visualised as a naked, voluptuous, blue-bodied female. 'Our connection is shrouded in mystery, even to me,' says Cortana. The original connection between the two AIs was in name only, an internal shorthand used by the development team, much like the foetal nickname given to an unborn baby by expectant parents. After an unexpectedly positive fan response to a leak of the code name, Cortana stuck, and a collection of more gender-neutral titles the team had been trialling were dropped. Subsequent meetings with the 343 team, the studio responsible for Halo, prompted Harrison to go back to her personality build and write in more confidence and self-assuredness, something that she admired in the Xbox version of Cortana. 'I thought that was a great characteristic and we'd like to represent that better than we had been doing. As a result of the meetings we ended up coming back and baking this into the DNA of the digital assistant speech pattern, her approach to jokes and her approach to chit chat, and we decided to dial up her sense of self possession,' Harrison explains. Unlike Halo's Cortana, the digital assistant is now just a pulsating blue circle and when asked if she is naked replies 'I'm ... code.'

The First Sound of the Future

Hatsune Miku, the Japanese AI pop star, is also female. The Japanese characters that make up her name translate as 'first', 'sound' and 'future', or together 'the first sound of the future'. Fans scream and wave their glowsticks at a Miku concert, drowning out the synthetic music with shouts of admiration for a holographic ghost. It is unclear whom they are cheering for – the animator, the programmer, the original voice actress, the motion-captured dancer or Hatsune herself, an assemblage of inputs just like Cortana. One of the people behind the scenes is Wataru Sasaki, lead engineer of Vocaloid, the software that powers Hatsune. In an interview in Japan, we discussed why the interface with AI is so often wrapped in a scantily clad female form or mediated by a woman's voice. When developing Hatsune Miku, the popular style in anime was Bishōjo, or 'beautiful girl', typically characterised by the deforming or exaggerating of the female figure. Sasaki believed the sexual symbols attached to this culture were excessive and unnecessary as such characters were appealing to a more intelligent and unique fan base. 'We excluded

We believe we need at least two languages through which to communicate with robots. Just voice, gesture, touch or appearance on their own is not enough but rather it is the combination of these modalities that makes our imagination more active.

Hiroshi Ishiguro,
University of Osaka,
Geminoid HI-4,
2013

Geminoid HI-4 is a tele-operated android modelled after its original creator Professor Hiroshi Ishiguro. The robot has 16 degrees of freedom, to mimic natural human movements.

typical Otaku features such as big boobs, red lustful cheeks and instead derived her features from the colouring and forms of the classic DX7 keyboard synthesiser.' Just like Cortana, Hatsune is a personified machine, a female form wrapped around a machine of computation.

Elsewhere in Japan, at the University of Osaka is Professor Hiroshi Ishiguro's Intelligent Robotics Laboratory, an uncanny valley of institutional offices inhabited by an unsettling collection of humanoid robots in varying states of undress. Here we meet one of Ishiguro's research collaborators, Dr Kohei Ogawa. Together they have been building multiple generations of androids based on their own idealised human forms to explore how people might emotionally connect with AI. In other countries there is a rising fear of robots taking over the world, stealing our jobs or making us slaves. In Japan, however, a country without a migrant labour population, the lab has been developing autonomous technologies to provide domestic support and infrastructures of care that we would typically assume to be only provided by another person. These technologies are embraced in a context where representations of robots in popular media have already enculturated the Japanese to some of the emerging conditions of the Post-Anthropocene. Ogawa believes this adaptation is also connected to their belief system and explains that 'Shintoism is a natural religion, with a god that is everywhere which means that they can see a mind or a soul in any kind of object.' He suggests that the current generation of digital assistants are too abstract: 'we believe we need at least two languages through which to communicate with robots. Just voice, gesture, touch or appearance





Crypton Future Media,
Hatsune Miku live in concert,
2007

Hatsune Miku performs concerts to her millions
of screaming fans as a digital hologram projected
on stage.

These consumer-facing
AIs currently repeat
biases and forms of
interaction which are
a legacy of human-to-
human relationships.

on their own is not enough but rather it is the combination of these modalities that makes our imagination more active.' The robots in Professor Ishiguro's lab reinforce the assumption that sufficiently intelligent machines should be representations of ourselves.

Machine-Native Interactions

Why do we presume that a device, a laptop or a toaster should have any binary gender in the first place? As Harrison notes, 'Not to disparage the concept of androgyny, but for us the female voice was just about specificity. In the early stages of trying to wrap our mind around the concept of what it is to communicate with a computer, these moments of specificity help give people something to acclimate to.' So how far away are we from a point where the crowd cheers not for a ponytailed, animated, female pop star styled after a classic synthesiser but for the DX7 synthesiser itself? Could we meaningfully engage with a gelatinous blob, or the monolith from Stanley Kubrick's iconic film *2001: A Space Odyssey* (1968), or do we intuitively require some form of personification for us to care? Perhaps the machine interactions of the Post-Anthropocene are not Scarlett Johansson speaking to us from a blinking vintage cigarette case (as the AI in Spike Jonze's 2013 film *Her*), but a voice that escapes the categories of male and female, a machine accent, like Spanish or French, that would be associated with a particular family of devices. Do we need to become fluent in C++ or the mechanics of machine vision in order to have a productive conversation with Cortana? How might we develop new forms of empathy or understanding with our technologies through alternative protocols of machine-native interactions?

These consumer-facing AIs currently repeat biases and forms of interaction which are a legacy of human-to-human relationships. The future of machine interactions is not natural conversations with latex-skinned, humanoid-shaped robots, but rather its complex relationships with driverless cars, mirrored black rectangles, giant infrastructural objects and planetary-scale logistical systems. Machines do not see or understand the world like we do, yet we insist on trying to push our interfaces with them through the forms of language and vision that we associate with ourselves. Perhaps what we should be doing is looking at how, at these early stages, we can be prototyping new modalities of communication. What are the most appropriate means through which we can have an interaction with an autonomous shipping crane loading containers into the hull of a ship at a Chinese mega-port? How do we pose a question to a warehouse filled with a million objects or talk to a city managing itself based on aggregated datasets from an infinite network of sensors, cameras and media feeds? My watch told me about a coffee machine it just met, my toaster wraps me in a warm embrace, the smart fridge smiles and the GPS asks us to turn back. The streets are lined with sensors and the city tells me a joke. 'I think like a computer thinks ... I think,' says Cortana. ▢

This interview is based on a telephone conversation between Liam Young in Los Angeles and Deborah Harrison in Seattle conducted in June 2018, and discussions in Japan with Dr Kohei Ogawa and Wataru Sasaki.