AI & SOCIETY
https://doi.org/10.1007/s00146-022-01433-y

CURMUDGEON CORNER

Will AI end privacy? How do we avoid an Orwellian future

Toby Walsh[1]

Received: 3 August 2021 / Accepted: 10 March 2022

Vince Cerf told the Federal Trade Commission in 2013 that "*privacy may actually be an anomaly*". He justified this bold claim with the observation that "*Privacy is something which has emerged out of the urban boom coming from the industrial revolution*."

There is a certain truth to his claims. Our conception of privacy has changed greatly across time and space. Even limited to a single culture, we see significant differences in privacy across history. Back in medieval England, for example, life was a lot less private. Most people could not afford to live in houses with separate living and bedrooms. The Industrial Revolution lifted the standard of living and made privacy more possible. However, privacy is more than having a room of one's own. It is also about having the privacy to discuss political change and dangerous thoughts in private. To vote anonymously. To practice the religion of our choosing. To live the lifestyle we wish. And many other privacies that we have grown to expect.

A vital question then is how do we stop AI from taking away many of these privacies, much as George Orwell predicted in "1984". In 2013, Abhishek Mehta, the CEO of Tresata noted "*Just like oil was a natural resource powering the last industrial revolution, data is going to be the natural resource for this industrial revolution. Data is the core asset, and the core lubricant, for not just the entire economic models built around every single industry vertical but also the socioeconomic models*."

Artificial Intelligence has been and will continue to be a major consumer of that data. Machine learning methods like Deep Learning currently require millions, if not billions, of training examples. If data is the new oil, machine learning is the refinery. The analogy between data and oil is one that should not be taken too literally. There are some fundamental differences between data and oil. Oil is a precious and limited resource. Data is neither very precious nor very scarce. Oil can only be used once. Data can be re-used without limit. Indeed, unlike oil, data can often be used to generate more data.

1 AI and privacy

But perhaps the biggest difference between oil and data is in ownership. Countries quickly claimed ownership of the oil beneath our feet and under our seas. But much data today is privately owned. Indeed, a few private data monopolies like Google and Facebook increasingly own most of our data. And whilst they are generating wealth from it, we the producers of that data are receiving little of the value. In addition, all this data is putting our privacy under increasing threat.

There appear to be few boundaries on the data that technology companies are willing to collect. Bruce Schneier, a privacy and security expert observed that "*Surveillance is the business model of the Internet*". Al Gore put it even more succinctly when he called it "t*he stalker economy*".

Why did Google think it should track your Android phone when you turned off your location and even removed your SIM card? Or Uber think it should track your location five minutes after your ride was over? Or Pokémon Go think it should have access to your entire Google account on iOS, including your email and browsing history? And George Orwell would likely have been surprised that it wasn't the government that was putting surveillance devices into our homes but consumers paying their own money to tech companies like Amazon to do so. Alexa, are you listening?

Of course, it's not just technology companies that are starting to invade people's privacy. States are also adopting new technologies to pry into our lives. The US State of Delaware has put "smart" cameras into police cruisers to detect vehicles carrying fugitives, abducted children or missing seniors. Such uses might be unproblematic but what happens when a state starts using the same technology to track political activists or refugees? In China, police in Zhengzhou

✉ Toby Walsh
t.walsh@unsw.edu.au

[1] School of Computer Science and Engineering, University of New South Wales, Sydney, Australia

Published online: 07 April 2022

have used face recognition glasses. These glasses can process 100,000 faces every second. You can't hide in a crowd of demonstrators when such technology is being used.

Some may argue that the battle for our digital privacy is already lost. We have already irrevocably given up too much of our private information to Facebook, Google, Amazon and others. But we will shortly give up our analog privacy too. The challenge here is that we are connecting ourselves to smartwatches, fitness monitors, and other devices that monitor our analog selves. Our geographical location. Our heartbeat. Our blood pressure. And soon many other vital signs.

The benefits our obvious. The DeepHeart machine learning app for the Apple Watch can detect atrial fibrillation, hypertension and sleep apnea. And it can even use your heartbeat to predict the onset of diabetes with a remarkable 85% accuracy. This is all part of the promise that Artificial Intelligence can make us healthier. But there are also many risks. What

2 AI and privacy

if your health fund increases your premiums every week you skip the gym? Or your employer fines you for working too slowly? Or an advertiser show you commercials to make your heart beat faster?

With our digital selves, we can lie. We can pretend to be someone that we are not. We can connect anonymously. But it is much harder to lie about your analog self. We have very little direct control of how fast our heartbeats or our pupils dilate. Imagine what a political party could do with access to everyone's heartbeat? And we are giving this analog data away to private companies. For example, when you send your saliva off to AncestryDNA for genetic testing, you have to agree that you grant them "*a royalty-free, worldwide, sublicensable, transferable license to host, transfer, process, analyze, distribute, and communicate your Genetic Information for the purposes of providing you products and services, conducting Ancestry's research and product development, enhancing Ancestry's user experience, and making and offering personalized products and services.*"

And if AncestryDNA happens to use your DNA to develop a cure for a rare genetic disease that you possess, you will probably have to pay for that cure. The AncestryDNA terms and conditions make it clear that "*you acquire no rights in any research or commercial products that may be developed by Ancestry using your Genetic Information*." Actually, it used to be worse. Before a media outcry, AncestryDNA claimed a "*perpetual*" royalty-free license. Once they had your data, there was no way for you to get it back. At least now, you can ask for them to delete your data and not use it anymore.

Analog data is not protected by any patient/doctor or patient privacy legislation. A company like FitBit or AncestryDNA can do pretty much what they like with it that advances their business. FitBit can work out who is having sex and try to sell them some viagra. AncestryDNA can determine that you are at risk of Alzheimer's and sell your details to a local care home.

As is the case in many other areas, Artificial Intelligence is not only part of the problem but is also a significant part of any cure. There are a number of ways in which AI can help preserve privacy. Perhaps one of the surest ways to keep hold of privacy is not to have your data leave your possession. We will soon have enough computing power on our devices that the computation can happen right there. Your smartphone will be smart enough to recognise your speech, understand your request, and act upon it without calling upon Google or any other service in the cloud. Your health monitor won't have to share your vital statistics with FitBit or anyone else. It will track your heartbeat and identify for itself when you need to see a doctor.

3 AI and privacy

We will also soon have AI privacy assistants sitting on all our devices. Their sole job will be to protect your privacy and defend your security. They will monitor all incoming and outgoing data, and intervene whenever your privacy or security is threatened. Other technologies will also contribute to safeguarding our privacy. For example, quantum cryptography will be commonplace, giving even greater security to our data. And technologies like differential privacy will be mature, letting us share data with others in society for the public good but without giving up our own privacy. Finally, if we can diversify those building the technology, we can hope for more inclusive and private outcomes that reflect better our different cultural values.

If we make the right choices today, privacy will not be a historical anomaly. It will be a technological given right.

Curmudgeon Corner Curmudgeon Corner is a short opinionated column on trends in technology, arts, science and society, commenting on issues of concern to the research community and wider society. Whilst the drive for super-human intelligence promotes potential benefits to wider society, it also raises deep concerns of existential risk, thereby highlighting the need for an ongoing conversation between technology and society. At the core of Curmudgeon concern is the question: What is it to be human in the age of the AI machine? - Editor.