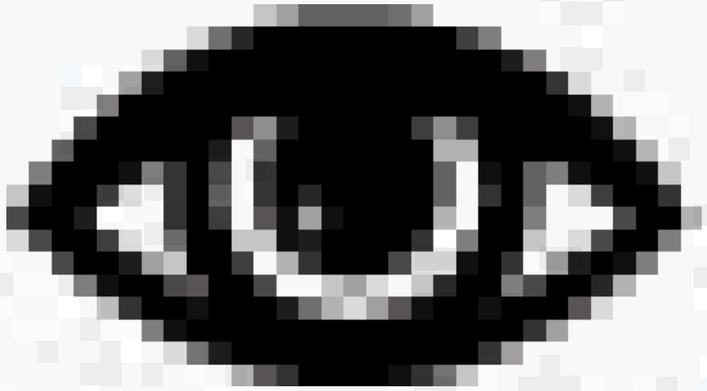


I always feel like somebody's watching me...

"What we're trying to do is develop processing frameworks that would allow this data to be useful and to be used, without the somewhat creepy feeling that you're constantly being watched"

What power can individuals have over their data when their every move online is being tracked? Researchers at the Cambridge Computer Laboratory are building new systems that shift the power back to individual users, and could make personal data faster to access, and at much lower cost.

It's a fact of modern life – with every click, every tweet, every Facebook like, we hand over information about ourselves to organisations who are desperate to know all of our secrets, in the hope that those secrets can be used to sell us something.

Companies have been collecting every possible scrap of information from their customers since long before the internet age, but with more powerful computers, cheaper storage and ubiquitous online use, the methods organisations use to gather information about people have become ever-more sophisticated. And sometimes those organisations know us better than our own families or friends.

For example, several years ago, the US retailer Target's data analysis had become so precise that they were able to determine, with astonishing accuracy, whether a woman was pregnant and how far along she was, based on her purchase of certain products. And in one particularly embarrassing incident, Target knew that a teenage girl was pregnant before her father did, much to her father's displeasure.

"What Target learned from that incident is that marketing too accurately can really make people squeamish," said Professor Jon Crowcroft of the University's Computer Laboratory. "But if they made their marketing a little less accurate by increasing the amount of privacy they give their customers, they found they can still retain or increase their customer base without making people feel as if they're being spied on."

Crowcroft's research is in the area of 'privacy by design' – systems that allow us to live in the digital world and protect our privacy at the same time. As the concept of the Internet of Things – internet-connected washing machines, toasters and televisions – becomes reality, Crowcroft insists that privacy by design is needed to address the massive power imbalance that occurs when our personal data is shared with, and sold by, corporations, governments and other organisations.

But privacy by design doesn't mean disconnecting from the online world and putting on a tinfoil hat – far from it. "There's already a lot of data stored about each and every one of us – the things we buy, the food we eat, the health issues we have – and for each of these market segments, there are perfectly legitimate uses for that data," said Crowcroft. "Collecting healthcare data is fantastically useful for tracking pandemics, preventative care, more efficient treatment, public health –those are all perfectly reasonable and positive uses for big data. At the same time, most sites gather information in order to target ads more accurately, and most people are actually ok with that. So the question then becomes, what is privacy by design?"

"What we're trying to do is develop processing frameworks that would allow this data to be useful and to be used, without the somewhat creepy feeling that you're constantly being watched," said Crowcroft's colleague Dr Richard Mortier.

The type of system which Crowcroft and Mortier envision is one in which the user has the scope to allow access to their data on a case-by-case basis, rather than it be harvested whether they like it or not: computations are performed where the data is gathered, and the results are pushed back to the organisation that wants the data.

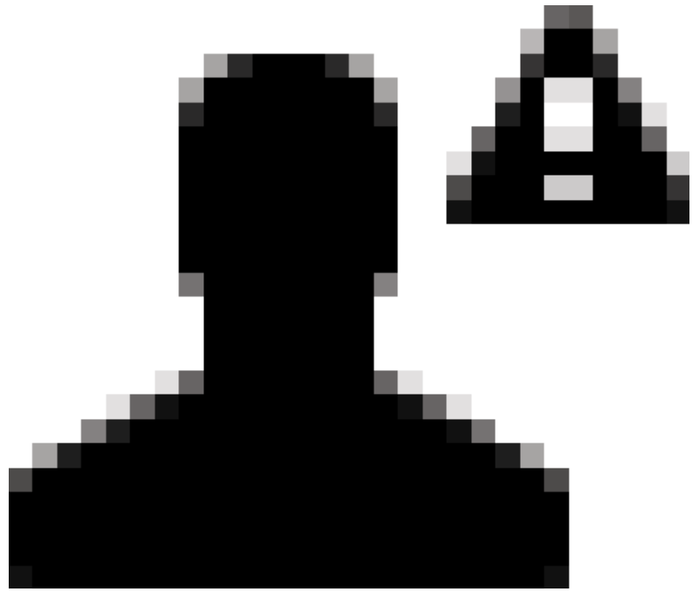
"We can change the big data problem completely by moving where the data is processed," said Mortier. "Rather than having systems where all of the data is gathered in some huge central location and processed, if you reconstruct the system so that the data is processed in the same place it's gathered, individuals would be able to take some of the control of their information back from corporations and surveillance organisations. Instead of one huge central processing node, we want to see billions of smaller nodes, which would make information quicker to access, and could potentially be stored at lower overall cost."

Crowcroft and Mortier have designed and partially built systems where a person's data stays local to them, and they can have the option to decide what is shared and with whom. For example, a patient can share their healthcare data with their GP, but the GP would have to get authorisation from the patient before sharing that data with a pharmaceutical company.

"People realise they're being marketed to, but I don't think they realise the scale of it – it really is a hidden menace," said Crowcroft. "The point is that we could build systems that could stop that completely, and re-enable it on the basis of a level playing field. We want to see systems where people have agency over their data, giving them the ability to allow or prevent certain types of access."

Contrary to what some people may assume about the nature of digital life, said Crowcroft, the vast majority of people highly value their own privacy. He points to the launch and then recall of Google Glass, a wearable computer worn like eyeglasses. "People started wearing these things into restaurants and other diners wouldn't put up with it, because they didn't want to be recorded while eating their lunch – it really creeped people out," he said. "And that's in a public space: imagine the same sort of thing happening in a private space. It's about the asymmetry and the idea that this is being done to you and you have no comeback. The problem with digital infrastructures is you don't see them, and to a certain extent companies depend on people not understanding them – we can build systems where there are mechanisms through which they can be understood."

Crowcroft and Mortier recognise that they'll never convince everyone to ditch cloud computing and switch to a centralised system. But that isn't their goal. "It takes a while to show that new ways of doing things can really work," said Crowcroft. "If these sorts of systems become a reasonably widely used alternative, it will go a long way towards keeping companies and cloud storage providers honest. The very small number of providers leads to the exploitation of the network effect, where they have a strong monopolistic position over a certain type of data. And monopolies are not good for economies. If a decentralised system is more ethical, enough people using it may incentivise the big providers to be more ethical too."



I **Left to right**
Professor Jon Crowcroft
Jon.Crowcroft@cl.cam.ac.uk
Dr Richard Mortier
richard.mortier@cl.cam.ac.uk
Computer Laboratory