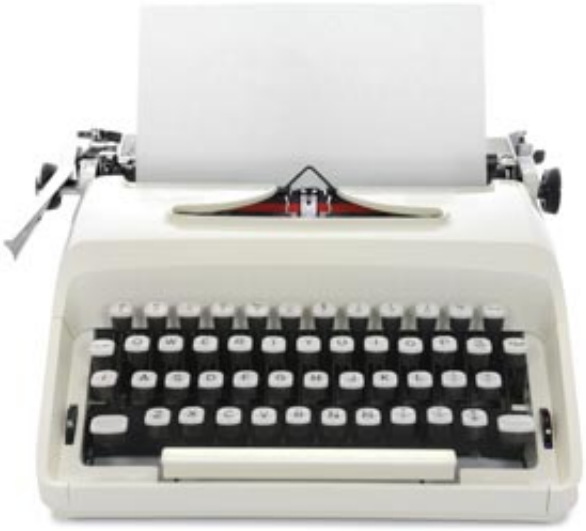
What are they doing with my data?

March 9 2016, by Lux Anantharaman



Before the lifting of the iron curtain, authorities in Eastern Europe required that anyone with a typewriter provide a sample typed page so any typed seditious material could be traced. Credit: Olivier Blondeau/Getty

You go shopping. You check-out at the cashier and are ready to pay. The cashier pulls out a camera and takes a picture of you, your bill and your credit card. You ask the cashier why. He tells you that the photo enables the supermarket to better profile customers like you—based on how you look, what mood you appear in, what clothes you wear, who you are with and what you buy. Demographic and behavioral information about you will be inferred from the information collected.

You tell him that you gave him your credit card number to complete the purchase, not to help the supermarket create a profile of you. You also ask him to delete the photo. He responds that this is company policy and that if you don't like it, you can take your business elsewhere. He adds that every other shop does the same—you don't have an option. Imagine being photographed every time you shop. Most of us would find this unacceptable, wouldn't we?

But something similar happens every time we shop on the Internet. We are constantly profiled based on our shopping habits. Companies do this so they can:

1. Recommend specific products and services
2. Price discriminate and extract as much value as possible from each customer
3. Influence you to change your buying patterns—during your pregnancy, for example.

In the 70s and 80s, in Eastern Europe, anyone owning a typewriter had to report to a police station once a year along with a typewriter and type a sheet of text. This enabled the police to track any typed material which they felt was objectionable or undesirable. Most of us today would find this unacceptable.

Credit: AI-generated image (disclaimer)

But how many people know that almost all modern day printers insert a barely visible set of yellow dots on every page printed, so it can be traced to you and the printer?

You and your significant other engage in some lovey-dovey online video chat, believing that no one else will ever see it. But what if the online service provider who you trusted to safeguard your data wasn't so secure after all and someone was able to hack the online service provider and extract millions of video chats? It sounds like something from a movie, but this is what happened in 2008, when Yahoo webcam feeds got hacked by British intelligence.

You are watching TV and a luxury hand-bag ad starts playing. You want to check out the hand-bag online using a shopping app on your phone.

When you open the app, you see a promotion for the very same hand-bag. You are wondering, how did my mobile shopping app know that I am interested in this hand-bag? Is my phone listening to my TV? If my phone can listen to my TV, can it also listen to conversations in my living room?

Something similar is becoming commonplace. Companies like SilverPush are trying to figure out all the different devices you own. Unknown to you, the ad on your TV emits an inaudible, high-frequency sound that your mobile app picks up. Now, your shopping app retailer knows that you were watching the TV ad and browsing for the same product at the same time. Cross-browser tracking is the latest obsession for Internet marketers.

In the above three cases, our personal data, ostensibly collected for one reason, was used for other purposes. In some cases, such as online shopping, we are vaguely aware that data is being collected, but in others we may not realize that our information is being collected (for example, inaudible sounds linking devices, invisible dots on laser printers). While for each device we all 'agree' to an end-user license (which almost no one reads), the data collected may be used (or misused) for purposes not explicitly made known to us.

Take e-commerce, for example. We provide our names, addresses and credit card information to complete a transaction, never realizing that the same information is used to create a digital profile. Even more worrying is the use of sophisticated big-data algorithms that, based on what we buy and where we travel, can infer our likes, dislikes and even our personalities. Researchers have found that based on Facebook likes, computers can judge our personalities better than friends, family, even our partners.

Our mental models of ownership tend to attribute value to physical

objects, whereas we don't typically view our personal data and information inferred from it as a commodity. Unlike tangible, physical objects, information can exist in multiple places at the same time—once it is released we may lose control over what happens to it.

But is it possible to know where our data is and how it is being used? Can we store data in 'capsules' in a secure, confidential manner and reveal it only to authorised applications and entities? Can the data capsule be cryptographically protected along with data-use policies? Can we design a system that allows authorized applications and entities to transparently access our data, only for specified time periods? How do we create policies that are simple to understand, but are capable of representing real-world richness? All these requirements become important in the world that we will soon be in, a world of Big Data and Internet of Things.

Building on my previous work on enterprise digital rights management platforms and secure end–to–end data protection, I am currently researching solutions for creating data capsules to make our presence on the public Internet more private. My objective is to create mechanisms, both technological and legal, to ensure that our personal data is only used for the originally intended purpose.

Provided by Agency for Science, Technology and Research (A*STAR), Singapore