Ard Su
Science & Technology

What Your Digital Footprint Says About You

Computer algorithms are becoming more adept at using our data to penetrate the deepest levels of our psyches.

By
Josie Cox '22BUS
|
Spring/Summer 2025
Ard Su

Sandra Matz, a computational social scientist at Columbia Business School and the author of *Mindmasters: The Data-Driven Science of Predicting and Changing Human Behavior*, discusses the complex relationships between our digital lives, our data, and our self-determination.

You argue that a person's digital footprint offers a window into every aspect of their life. What do you mean by that, and how are these footprints created?

Digital footprints are the traces we leave as we interact with technology. We generate millions of them every day. These include the traces that we intentionally put out there: things we post on social media, people we follow on LinkedIn, our Spotify playlists, and our Pinterest pages. Such interactions announce who we are, and social scientists call these "identity claims."

But there are other digital footprints too: they're the traces we leave without really thinking about it. These are called "behavioral residue." They're created as we interact with technology — when we swipe our credit cards, search on Google, shop online, accept cookies, or use smartphone apps that generate GPS data. Our phones

can collect data on our location pretty much 24/7 because we take them with us everywhere we go. Each of those data points represents a little puzzle piece of who we are, and once you put the pieces together, they provide intimate insight into our lives.

You say that all these data traces make us vulnerable to psychological targeting. What does that mean, and how can it happen?

On a basic level, the way I think about psychological targeting is that it describes the ability that algorithms have to decode our minds. The data in our digital footprints can be used to make inferences about who we are and how we feel, think, and behave. This includes not only inferences about where we shop or eat or travel but also inferences about deeper psychological dispositions: our personality, our values, our political ideology, our sexual orientation.

We can use an AI-powered predictive model to take individual pieces of raw data about us and put them all together to create a picture of a person — our desires, needs, preferences, motivations, dreams, hopes, and fears.

But the second stage of psychological targeting — which to me is the more interesting part — is that once I understand who you are, I can use that data to potentially change the course of your life. I can try to influence you and decide what information I want to give you. That might be a particular type of news or particular products and services.

It’s easy to see why advertising agencies might be interested in psychological targeting — for example, a mortgage company might want to show its ads to someone who spends a lot of time looking at real-estate listings. That seems relatively benign, but you argue that psychological targeting can be a powerful weapon. How so?

If someone’s personal data gets into the wrong hands, it can open the door to a whole new level of discrimination. In some countries, information about your spending habits, your mental health, your sexual orientation, or your political affiliations can determine whether you’re able to get a loan or a particular job. Psychological targeting can also be used against you in other ways. It can be used in an attempt to influence critical choices such as who you vote for. It can be used to fuel hatred or fear. Once I understand what terrifies you, I can play into that — even without your being aware of it.

You've demonstrated through your research that psychological targeting can increase polarization and create and fortify echo chambers.

Echo chambers are part of human nature. We all surround ourselves with people who are somewhat similar to us, and it's comforting when our worldview is reflected in another person. Having our beliefs affirmed and our identity affirmed is something that humans crave, and technology amplifies these natural tendencies. So if you have an algorithm that knows exactly what you believe in, what your past preferences have been, and what your past behaviors have been, then that technology can reinforce all of those preferences. It can understand something about your past and project that into the future.

I think there are two reasons why this is potentially dangerous. First, if my reality is completely different from yours, there's no way for us to come together and talk about our shared experience of the news, for example. If I don't even see what you see, then we won't have a shared political context and we won't have a shared cultural context.

AI has become so good at personalizing content. I could, for example, ask it to create a James Bond movie that's customized perfectly to my interests. Maybe it features the actors I love, maybe it has a storyline I find compelling, and maybe watching that movie feels great. But when we watch movies, we often want to discuss them with someone else. It's a shared reference point. But if my movie is totally different from yours, we lose that ability to create a cultural moment and have interesting conversations about it. So on an individual level, I'm worried that it might make us totally boring. What's more, if we're in our own echo chambers and the only thing that we see is what we believe in anyway, we are going to become more and more narrow-minded, and that's going to affect how we show up in the world. We don't talk about this enough, and it's not just a hypothetical. It's something that's already happening.



Sandra Matz (Columbia Business School)

Are there also ways psychological targeting can be used for good?

Yes. It could help make us more open-minded by giving us access to other points of view. Let’s go back to the echo-chamber example. Because of the algorithms that understand who we are, there are many entities — Google and Facebook, for example — that can conjure up the realities of many different people across the US and across the world.

My favorite example is that of a fifty-year-old Republican farmer in Ohio. I have no idea what that farmer’s day-to-day experience looks like. And I also don’t fully understand how he forms his opinions on immigration, for example, or on abortion. I’ve never seen his daily reality play out. But the same algorithm that Facebook uses to keep me in my own echo chamber could potentially allow me to hop into someone else’s echo chamber. Facebook knows what the fifty-year-old Republican farmer in Ohio sees, so Facebook could theoretically give me the option of going into explorer mode and seeing that reality too.

We could also use psychological targeting to help people reach their goals, to be healthier, even to save money. We see a lot in the news about how psychological targeting is being used in marketing to sell people more of what companies want them to buy. But we've done research on flipping that concept on its head, and it turns out that if we understand what motivates people to buy, we can also understand what motivates them to do the opposite: to not buy.

The idea of echo-chamber hopping is compelling, but how feasible is it? It's a big ask to convince someone to put themselves in a situation that might be uncomfortable or even confrontational.

You're right. My intuition is that people would not use it very often, because it's more comfortable and convenient to see what we like to see, but I like the idea of having the option. Right now there's no way for me to see what might have made someone support reversing the *Roe v. Wade* Supreme Court decision, for example. There's no way for me to understand or explore their point of view. So for me, it's not necessarily a question of whether many people will actually use it, it's more that I want to have the option.

Of course, there is a world in which this idea backfires. If I hop into someone else's echo chamber and what I see just really appalls me, it might make me dig in my heels even deeper. Instead of creating empathy, it might create disgust. So I've been thinking a lot about how you might be able to design the feature to make it about the person and not about the specific topic; how might you be able to create a deeper level of empathy. The idea of letting people walk a mile in someone else's shoes, if you will.

In your new book, *Mindmasters*, you give many examples of how access to information can be extremely dangerous, including the Nazi Party's use of certain records to identify Jews during the Holocaust. One way of preventing history from repeating itself might be to strengthen legal protections for such records. How might this be done?

This question is certainly on people's minds, since we give so much personal information to governments and corporations every day, and data on religious or political affiliations, income, health, sexuality, or other demographics could potentially be used against you.

When I think of potential regulation against this backdrop, I think a lot about the concept of privacy by design. In other words, how do you create processes that make it easier for people to not have their data out there? The way it currently works is that even if you wanted to protect your data, it's an impossible job. In the terms and conditions of all the services and products we're using, the default is usually that we have to actively opt out if we don't want our data used. And in some cases it's not even a choice: you either agree to all of the tracking or you can't use that product or service at all. If that's the choice we have to make, we're probably going to go with convenience and personalization and the ability to use a particular product. Therefore, one potential idea for regulation would be to require changing that default.

The more interesting solution for me, though, is a technical one called "federated learning," which is a collaborative way for companies to train AI models without having unfettered access to a person's raw data. It lets you keep your data on a personal device like a smartphone and makes use of the fact that your phone is effectively an extremely powerful computer in your pocket. If you take Apple, for example, when the company trains Siri, it doesn't collect all your speech data and send it to a central server. Instead, it sends the intelligence to your phone, which allows the model to work locally to learn and get better. And then anonymized information can still be sent back to Apple so that everyone can benefit from the improvements, but personal data is also somewhat protected.

To my mind, mandating this sort of model is a much more promising solution, because it doesn't require the individual to trust a company to not collect data in the first place. Because even if you trust a company today, who knows what might happen tomorrow.

You say that a lack of tech literacy is one reason so many of us are sharing our personal data without fully appreciating the repercussions. What can we do to enhance this sort of literacy?

I think we absolutely do need tech literacy, and we're starting to see an acknowledgment of this in schools. Even just a basic understanding of data and coding and an understanding of why data is intimate and the risks of sharing your personal information is very important.

But recently I've become a lot less optimistic when it comes to our ability to use data literacy to solve problems. Even if you keep up with technology, even if you fully understand the potential for abuse, it's a full-time job.

If you had to go through all the terms and conditions of every app and scrutinize your every digital interaction and consider exactly what data might be collected, you would have to say goodbye to all your free time, because that's a full-time job.

So putting this burden on the user and saying that they can manage it by themselves — that's an illusion. Do we need data literacy? Yes. It's absolutely necessary. But it's not sufficient.

You suggest that breaking up large tech companies and enforcing stricter antitrust laws would be one way to stop companies from amassing too much power and abusing it. How optimistic are you that this could happen?

I've become more optimistic over the last couple of years. People like Columbia law professor Tim Wu and NYU marketing professor and podcast host Scott Galloway have been advocating for antitrust legislation for a long time, but now policymakers are also waking up to the reality of how important this might be.

In the past, the argument has been that if you break up big tech companies, they won't be able to do what they're good at, because they won't have all the data they need to be able to do it. The argument is that they'll lose their ability to innovate. But we just don't know if that's true. One of the things we've seen is that placing constraints on companies just encourages them to innovate in a different way.

How do you personally protect your online privacy?

I get asked this question a lot! Obviously, I care about this topic. I think about it constantly. I have a fairly good understanding of where data gets collected, but even I don't manage it properly. I try to be a bit more mindful than the average person about managing cookies. But there are still so many times when, say, my baby is crying and I need to quickly find something online, so I just hit "accept all."

I might be a little more mindful with my phone, because I think the phone is so intrusive. When I download an app, I do go through the permissions, because I don't necessarily want my weather app to tap into my picture gallery and my microphone. But other than that, I think my own experience and my own failures to protect my

privacy are a really good example of why we need something that protects access to our data in a more systemic way. We can’t just place the burden on the individual.

Read more from ***Josie Cox '22BUS***



Guide to school abbreviations

All categories >

Read more from ***Josie Cox '22BUS***