Researchers Have Few Guidelines When It Comes To Using Your Data Ethically

Jessica Baron Contributor ⓘ
Consumer Tech
I write about the ethics of science and technology.

GETTY

Most people don't think about the amount of data they generate over the course of a normal day. But, in fact, you transmit information every time you pick up your phone, get on your computer, strap on your smartwatch

play a video game, or use your credit card. Data about you is gathered and stored by hospitals, insurance companies, schools, utility companies, courts, government agencies, banks, and more.

While this data certainly can and has been used for sinister purposes, it's most often used by researchers (putting marketers to the side) to get insight into human behavior, usually to make our lives easier. But unfortunately, there are very few guidelines when it comes to the responsible use of this data.

In a recent paper from the journal *Ethics and Information Technology*, a interdisciplinary team led by researchers from the University of Melbourne has given us 5 issues to consider when thinking about how to improve ethical data collection, storage, and research.

1. **Consent**

Most of us don't even seem aware of the data we're generating when we g online, making consent a difficult issue (at least in some cases). Are all of your public tweets fair game for researchers to collect and store and use a data points? Do you know if your location information is enabled on socia media? If you don't know the answers, are you tacitly giving researchers permission to collect this information?

Technically, we all click on legal agreements when we use websites and apps that collect our data, but few of us read them and even fewer understand them. Even back in 2017, Deloitte found that 91% of people ir the U.S. consent to terms and service conditions without reading them (97% in the 18-34 age group). To be fair, even back in 2008 law scholars estimated that it would take the average American 244 hours a year to rea the privacy policies of all websites he or she visits.

If researchers can't be sure participants understand how their data will b collected, stored, and used, how can one achieve informed consent?

Informed consent is not simply something we can toss to the side just because it's difficult to attain with new types of data. In the U.S., the Federal Policy for the Protection of Human Subjects or the "Common Rule," published in 1991 outlines the basic provisions for informed consent and other elements of ethical research and exists because mankind has a bad track record when it comes to heinous experimentatic on humans. But it doesn't fully address issues of data collection and use i academic or government research. As a result, researchers sometimes fee those rules about consent don't apply to them.

The Australian researchers also discuss the difficulty of tracking consent when data is repurposed and used for new and unanticipated analyses. W don't always know where our data will end up and how it will be used, bu we can do our best to give it away selectively. However, in the end, the onus is on researchers.

As Dr. Assunta Hunter, a co-author on the paper, told me in an e-mail:

> "Researchers have the ultimate responsibility for ensuring that the dat they collect as part of any research project is safely stored, securely managed and ethically used. This is part of the researcher's contract with participants to carry out research in an ethical manner. The oversight and responsibility for data collected must be a part of the planning, the implementation and the writing up of the research.

Our digital ignorance as citizens may be helping make the case for loosening regulations on informed consent or insisting it simply doesn't apply to online data. But instead of spending hundreds of hours trying to

understand privacy policies (most of which we might not agree to anyway if we understood them), we can at least educate ourselves about our privacy options (click on that privacy tab on social media and turn everything off!) and register our discontent with unclear privacy policies. We should not only understand what we're consenting to, but demand to opt out later, either turning off the data tap or removing our information from repositories where it's stored.

2. **Privacy and confidentiality**

Privacy and confidentiality are not the same things. Privacy is regulated b law (and is not an inalienable right) while confidentiality is a professiona ethical obligation.

While we should expect researchers to protect both, there are cases in which the common good can sometimes override what we feel are our rights to keep our information to ourselves. For example, public health ha always relied on surveillance and data collection including age, location, travel, ethnicity, and more in order to track and halt the spread of disease There are strict rules for its use, of course – as there should be for digital data – but the point is that there are plenty of ways to use data collection for the common good that do infringe on privacy and confidentiality. But we should expect researchers to have very good reasons for collecting and using our data.

It's important to learn how those who run our web browsers, store our e-mails and photos, maintain social media networks, collect information from our apps, etc. intend to keep it private, as well as how they intend to follow (or lobby against) new laws that increase regulation of its collectio and use, and to what extent they hold themselves to an ethical standard t maintain our confidentiality.

Dr. Hunter told me that she and her research team agree that participant in research "should be confident" of at least three things:

1. That "the anonymity and the confidentiality of the information/data will be maintained in the way that the research was conducted, published and used."
2. That "they would be notified of any data breaches/ hacking/loss of data that occurred in datasets/research projects where their information was recorded."
3. That "a data governance committee and/or data security manager would be in charge of overseeing the security and storage of their data."

Problems also arise here when researchers don't invest in keeping data secure, restrict it to authorized users, and vet those users carefully. Once our data has been used by those who collected it, we must rely on them to keep it safe. That's a tall order and a lifelong commitment. That's one of the reasons privacy laws are needed to mandate the storage or destruction of data once researchers are done with it.

3. **Ownership and authorship**

Because data is valuable, it's very rarely destroyed. Privacy and confidentiality *should* prevent researchers from selling it or passing it on without our permission, but that requires us to have a conversation about who really owns data to begin with. Claiming our data always belongs to us actually gives researchers an easy out when it comes to abandoning it once they're done with their project.

The Australian researchers bring up a list of difficult questions we must address: “Are data owned by participants, research funding bodies, the principal researcher, members of research teams who collected the data, the guardians of data, or data storage services? In the case of digital data emitted by a device, is the person who is using the device considered its author or is the owner the person who has collected the data for research purposes?”

It’s important to be clear on this because ownership of data requires researchers to be accountable for it – to uphold their ethical obligation to confidentiality and their (potential) legal obligations to maintain our privacy. This doesn’t necessarily help solve long-term problems of how to protect and maintain data, but it does provide a trail of ownership to follow when data is misused.

4. **Governance and custodianship**

“Researcher” is a vague term and those who do research exist in academi industry, government, and other settings, each with their own guidelines. In addition, much modern research is international in scope.

How data is governed and who has “custody” of it for the purposes of assessing accountability adds yet another layer of complexity to the discussion. Who gets a say in how data is repurposed or shared? Where will it be housed? What responsibilities are transferred when data acquired by a new entity? And, returning to our first issue, how do we achieve informed consent when data is transferred?

Since data is a useful commodity and research tool, we also need to consider who may have a right to use it for their research. Can we reasonably expect every researcher to re-collect data for each new project

If not, what does a data repository look like and, again, who governs its use? A gatekeeper must maintain security, but under what conditions should they allow access to other researchers? How can it be shared safel and who is responsible for misused or lost data in this case? Will smaller entities with good intentions but low security budgets lose out on accessing vital research data?

5. **Data-sharing**

Questions about the need (and even responsibility) to share managing, preserve, and secure data are, of course, numerous and complex. The complexity is compounded when data is controlled by a commercial entit who may have no obligation to share the fruits of their labor. Even academics (who are in many ways obligated to share their research) may not always be willing to share what is increasingly a valuable commodity with competitors.

Open data portals are being developed in some fields, such as genomics, t ensure access to data by researchers as well as the public. While this is largely de-anonymized, it's worth appreciating the public service being done by those housing it.

It's clear we need specific guidance at the industry, national, and international level when it comes to the ethical issues posed by big data in order to minimize the risk to people. The ethics committees that already exist - in universities, for example - do not have the tools necessary to make informed decisions when assessing research proposals on processir new and repurposed data sets.

The Melbourne team notes that while there are few accepted best practice in this area currently, researchers and ethics committees need to talk openly about how and why data is being collected, what safeguards are in place, and create internationally relevant guidelines with the help of a diverse group of individuals. Hunter elaborated in her e-mail:

> “ The widespread use of international guidelines is desirable because research teams and projects are increasingly carried out by teams working in different countries. Participants in web-based research may live in many different countries. Additionally, data may be stored in off-shore storage facilities and the movement of data across national boundaries in this circumstance is increasingly common.

It’s no easy task, but outlining some of the issues is the first step in tackling this complex and ever-growing set of dilemmas.

Jessica Baron

I earned a Ph.D. in History and Philosophy of Science in 2013 and work as a writer ar consultant on a wide range of topics including the ethics of science, techno...
Read More