Article

14 minute read • 22 March 2023

A new narrative for digital data

What Indigenous cultures can teach us about digital privacy

Peter Evans-Greenwood
Australia

Prof. Deen Sanders OAM
Australia

Rob Hanson
Australia

Finding a better metaphor for our dealings with data

We understand the world through stories. They both define and describe the paradigm of our world, enabling us to place ourselves in the narrative and use metaphors to link complex concepts and make comparisons between new and

familiar experiences. Stories can be powerful tools for finding our way through new worlds and new information, helping us leverage hard-won knowledge and experience from one challenge into the next. However, they can also cause problems when a new challenge fails to follow the narrative rules of the existing story. Such are the challenges with data privacy.

Western societies rely on the metaphor that data is private property and consequently is exchangeable, tradeable, and ownable. It projects ownership concepts associated with physical objects—such as protection and use—onto intangible digital data. Framing the story of data in this way invariably leads to discussions on digital privacy based on a narrative of "rights." When is it permissible to collect data associated with an individual? Who has the right to say how a particular data set can be used? How are these rights enforced? Under what circumstances must rights be ceded to others?

Acknowledgement of country

My people (the Aboriginal and Torres Strait Islander people of Australia) are the world's oldest continuous culture. We have lived through time measured on a geological scale, resolving the complex problems of property, privacy, markets, and governance in ways that were encoded into our roles as a custodial people for nature (country), its animal, plant, and people communities.

At moments of uncertainty and complexity like the ones we face in the world now, the answer lies—as it always has—in a better understanding of our relationship with the world and ourselves. As ever, it is the knowledge and responsibility of Indigenous people, as custodians of country, that can best help everyone come into a proper relationship with country, with *Nayiri Barray*—our Mother.

The authors of this paper (Indigenous and non-Indigenous) come from all lands, and we pay our respect to the elders (past and present) of the lands from which contributions to this paper have emerged, and the land on which you are reading this article now.

I acknowledge that all of us are only here today because of the sacrifice and curatorial responsibility our elders carried and continue to carry for these places. I also want to acknowledge any Indigenous brothers, sisters, and elders who read this report or who are asked to give their leadership and knowledge, because we share responsibility across the globe for building a future, healthier, shared system.

Marrungbu
Thank you.

Professor Deen Sanders, OAM
Proud *Worimi* man and Deloitte partner

A property-based approach requires us to apply tools developed to manage property rights to digital privacy. By default, it makes an individual—as the owner of the data—arbiter of how *their* data should (or shouldn't) be used. We're drawn to this approach because it fits our familiar narrative of market exchange and aligns with modern tenets of individual self-determinacy and autonomy. The problem is that this approach isn't working.

People cannot begin to fathom the ocean of data, both metaphorically and literally, that is being collected and shared about them. We are saturated with data, and the world is saturated with data about us. But what if none of that data is true? How can the policy decisions of government, corporations, and individuals be correct if the data relied on is not the whole truth but merely a one small glimpse of a larger picture? Companies and governments are operating in deep waters. Alone in that ocean, individuals are powerless to swim against the currents and whirlpools of license agreements, privacy policies, and sometimes, outright fraud.

Treating data as property is not working. But there is another narrative that might open a more productive path to understanding data privacy—one we find in Indigenous Australian cultures, the world's oldest continuous culture. These metaphors revolve around responsibility and relationships instead of

rights and can provide a new and more useful understanding of data's role in society and our dealings with it.

In Indigenous Australian culture, it's the relationships among actors, rather than the actors themselves, that define the world. All things in the world—people, property, and the country itself—are considered actors with a role to play. Social norms are framed in terms of responsibilities—actors' obligations to each other within their web of relationships. All actors have responsibilities, and so all things are designed to support all other things in meeting their responsibilities. One's agency—the capacity to act—is determined by the intersection of these responsibilities.

As an example, consider Western sustainability challenges in light of this metaphor. Governments might sign a right to clean water into law, for example, but it is the actions of individuals (and individuals in organizations) that pollute local water sources like rivers. Clean water is the outcome of each member of the community upholding their responsibilities—disposing of waste appropriately, maintaining catchment areas, etc. One's agency, in relation to the water source, is the directly tied to meeting one's responsibilities.

But individuals and organizations aren't the only actors in the Indigenous narrative. For the system to work coherently (and sustainably), the actors in the web of relationships must encompass both the animate and the inanimate—the features of the landscape and the flora and fauna that inhabit it, just as much as the people and communities that populate it. As one of our authors reminds us, Aboriginal culture does not hold humans to be any more important than any other aspect of country—we're no more important than the kangaroo, the tree, or the river. All have a role in the community, and all are seen to have reciprocal responsibilities with the other actors with which they interact.

An incomplete picture

Did you notice the beautiful image at the top of this story? Perhaps you assumed that it is authentic, drawing as it does on a color palette of natural ochres and embodying geometric and natural symbols that are frequently found in Aboriginal art.

But it is not the work of an Aboriginal artist. It's not even the work of a human being. The image was created by a generative AI platform, responding to the prompt: "Imagine a color image of the data world in the style of Aboriginal art."

It looks like it might be authentic, and it's certainly pretty—but what does it really mean, and can it have any meaning when it is devoid of context and knowledge? Aboriginal art is not just image: it holds knowledge and is part of a story that has been orally transmitted for thousands of generations in specific relationship to place, kin, and culture. It is based less on interpretation of data and more as a form of conveying data that has been wrapped in endless layers of history and complex, relational knowledge. How can meaning be conveyed when the platform simply assembled known "data" about Aboriginal art, and stripped it of context, knowledge, and humanity? Meaning and knowledge are lost between the original data and the presented art. It's even more problematic for knowledge transmission that relies on the infallibility of story. How do we know what is artificially added by the generative system in order to present something that looks "artistic?"

Data about (or on) individual human beings is a record of facets of us, exclusively in relation to the system seeking the data. But human beings are meaning-making entities, and so we dutifully seek to ascribe new meaning to those elements and invent a picture (or knowledge) about the person from that data. Indigenous principles of data encode the data in the story. They are indivisible. They are relational. They lose power, meaning, and purpose the minute they are abstracted.

In the future, the blurry distinctions between true knowledge and artificially constructed knowledge will demand a new understanding of our relationship with data. The secret will lie, as it has always, in the way human beings construct knowledge and data dream.

The river, for example, has responsibilities to flow, to nourish and water country, and it is our responsibility to support the river in that function. We fail in our responsibility when we interfere with the river meeting its responsibility. This Indigenous concept of relationship and responsibility was the motivation for New Zealand granting legal personhood to the Whanganui River, the culmination of a 160-year campaign by the Maori, the indigenous[2] Polynesian people of New Zealand, to obtain legal protection for the river.

Though this is necessarily an oversimplification of Indigenous cultural perspectives, it provides food for thought as an alternative metaphor for understanding the complicated nature of data privacy in the modern world. In examining why our existing rights-based approach to data as property is fundamentally flawed, and then applying the Indigenous metaphor to better understand the relationships and responsibilities various actors bear in the realm of data privacy, we can begin to make progress in developing a more productive path forward.

Why data resists being owned

Before exploring data privacy in light of this new metaphor, it's useful to examine in more detail why the current metaphor of data as property is problematic.

A fundamental difficulty is that the digital is not physical: Data has different attributes and affordances from things that can be seen, touched, and held. It's ephemeral—something that we can only interact with via digital media and tools, unlike a physical object which we can interact with directly. The social convention of ownership of a physical asset—*this* individual owns *that* object—is unsupported by the nature of the digital world. Nor can we assume that a single set of social conventions will[3] be universally applied or enforced, varying as they do by culture and context.

Another meaningful way that digital data differs from physical objects is that it can be copied, transmitted, and shared, instantaneously and effortlessly, in a way that a physical object cannot. Even data in physical form (a printout, for example) is not equivalent to digital data, as it requires much more effort to

obtain and use. Searching a public property register, for instance, once required a trip to the local records office to spend time with the card catalog. Today, this can be done online from a distance. These records were always publicly available as physical objects, but now, as digital records, they can be easily found, collated, and (re)published.

Because digital data is so easy to copy and share—not to mention collect—it is prohibitively difficult for any entity to control who collects and copies it, where it's sent, and how it's used. Once released, the genie is impossible to put back. This already calls into question the utility of property as a concept for data, where ownership of property confers the right to control. The fact that digital data can be collected, copied, shared, and used without the putative owner's awareness only compounds the problem. Moreover, there is no simple one-to-one relationship between data, the individual, and the facts inferred from the data, as there is with physical property.[4] A datum might yield little privacy-compromising information on its own but reveal much more when combined with other data; for example, when a pattern across a data set provides insight into an individual's behavior.[5] On the other hand, a single datum might allow us to infer a range of facts relating to a range of individuals. A DNA sequence, for example, can be used to identify an individual, but it also enables us to identify others who are related to that individual: siblings, forebears, and descendants. The same DNA data can also be used to determine the propensity for various diseases in each of these groups. When data can implicate multiple individuals, who should control the use of that data? Should an individual be allowed to submit their own DNA to a service that conducts population research when that DNA data can also be used to identify others?

It's in the face of these attributes that the analogy between digital data and physical property breaks down. The digital realm has its own nature, and this nature is different to that of the physical realm. Applying traditional concepts of ownership to data doesn't work because data, by its very nature, resists being owned.[6]

Rethinking our relationship with data

If data isn't property, then what is it? If we look to the Indigenous Australian metaphor to provide a new and more useful understanding, the assignment of agency and responsibility to the inanimate and nonhuman holds a lesson for Western cultures in their conceptualization of data. In a very real sense, digital data has a life of its own. It's agentic in the sense that it acts on us and sometimes instead of us; think of census data driving government policy, for instance. When someone uses data about us as a way of pretending to know, understand, or serve us, it's the data they are knowing, understanding, and serving—not us.

This is further complicated by the fact that no representation can be complete. Our digital identity is not a single avatar—an assemblage of the discrete facts and figures that we leave littered behind us as we move about the world, an assemblage that somewhat resembles us. It's a whole field of avatars, each representing a different collection of facts and figures and so providing differing and possibly conflicting representations, but all thinking that they represent our one true self. Consumers of these avatars often confuse the avatar with the individual, assuming the avatar they see is a true representation rather than an assembled and biased image.

Like all actors in Indigenous Australian culture, data is also relational. Our relationships with people, place, work, government, markets, and the environment are unique to every human being. Data's meaning and value, and so our data privacy, is defined within the context of a relationship. Data that is private, personal data in the hands of an individual, may be marketable property in the hands of a data aggregator and a set of design inputs in the hands of a company using data to create a product or sell a service. It changes shape and function as it moves through different systems which have different relationships with both the individual and each other.

The assumption that data is property, and therefore subject to rights, has led to a focus on data users and holders (those who host the data, control the systems, and monetize or act on data as property) while underemphasizing the fact that one's relationships in the digital world are multidirectional. If organizations and people are to have rights in this landscape—permission to access or withhold data—and so have agency, they must first accept certain

responsibilities. Thinking of data as an actor in this complicated web of relationships opens the door to understanding its roles and responsibilities toward us, and our responsibilities toward it in turn. It prompts us to ask questions that may lead to fruitful dialogue instead of the deadlock the ownership debate has reached.

Consider identity theft as an example.[7] The term itself is an oxymoron as identity is not an object that can be lost or stolen, i.e., identity is neither a possession nor is it property. The crime that should concern us is *fraud*, where another actor impersonates an individual for the purpose of committing a crime.

Fraud is an ancient crime, with documented examples dating back to as early as 300 B.C. Digital media is the engine room of modern fraud, allowing a criminal to fraudulently present digitally accessed and copied information to an online service to convince the service to provide a loan, process a (false) tax return, or even request (undeserved) disaster relief.[8] There are two questions to concern us here. First is how the fraudster obtained copies of someone's personal data. The answer to this question might be insufficient security by some organization (or even the individual), or social engineering whereby an employee of some organization mistakenly provides the information. The second question is how the criminal used the personal data to misrepresent themselves and commit the fraud. It's surprisingly easy to, for example, obtain a credit card with only a few personal details. Framing this as a property crime, as identity theft, leads us to focus on making the data harder to steal.[9] We rarely consider making it more difficult to commit fraud, to misrepresent oneself using others' personal data, however.

As mentioned earlier, a collection of seemingly personal data does not provide one with a full understanding of the person it is claimed to represent. At best, it is a doppelgänger—an incomplete and biologically unrelated lookalike of a person. At worst, it might be a melange of data from multiple sources, each associated with different people, fused into a Frankenstein that represents everyone and no one. Fraud doesn't involve the individuals whose personal information was used to conduct the crime, and so the solution should not involve them. The solution to the problem of identity theft is to review the

responsibilities of organizations using personal data. A fraudulent loan, for example, should be the responsibility of the bank that provided the loan, not the individual who didn't request it.

Finding value in data responsibility

If we think of data as an independent actor in the indigenous metaphor, with its own complex relationships to navigate, we must also consider the scope of responsibility when it comes to data privacy—both for the data itself and the individuals and organizations using it.

Digital privacy is a complex issue, leading to complex regulatory and legislation solutions, such as General Data Protection Regulation (GDPR). Firms struggle as compliance costs (the cost of securing and reporting on valuable data) grow. Despite all this effort, breaches, where private data is hacked, are far too common. Moreover, the consequences of these hacks fall disproportionality on the individuals whose data was leaked.[10]

Discerning between legitimate and illegitimate uses of private data is also difficult, especially when the owner of the data (the human or nonhuman source) is rarely part of the transaction. Using data obtained via hacking is clearly an illegitimate use. But what about data firms harvesting information from the environment to create synthetic profiles for individuals—data which may or may not be considered personal information? In one recent example, data brokers created "sucker lists" of vulnerable people, such as the elderly, which they then sold to scammers. The lists were then improved by noting which people responded to misleading print solicitation.[11] This resulted in cascading fraud, where the same individual was approached repeatedly.

Some of these uses might even appear legitimate at first glance—using driver behavior data gleaned from public cameras and sensors, for example, to set individual insurance premiums—but in fact live in the broad grey area between legitimate and illegitimate. An initiative to align premiums with risk and so reward safe behaviors can easily tip over into a biased approach that hurts the less fortunate, such as when low-income individuals' car insurance premiums rise as they're forced to live in low-rent but high-crime

neighborhoods, or the vehicles that they can afford to purchase have inherently fewer safety features.

Regulating how data is collected and processed[12] is insufficient. In returning to the issue of fraud, for example, authentication procedures have favored convenience over accuracy. The resulting situation is one where it's easy for a fraudster to misrepresent themselves and obtain loans, tax returns, etc. These processes often highlight their convenience as a way of speeding up transactions but could simply be seen as a mechanism to speed up data acquisition. The truth of most of these data acquisition systems (e.g., convenience apps) is that consumers choose to participate on the assumption that their data is irrelevant, or their digital presence is a benign one. Authentication systems need to be strengthened, the scraping purpose of convenience apps need to be stopped, and consumer responsibility for their digital identity needs to be made clear, perhaps using a different metaphor to the one that gave rise to the theft in the first place. Ultimately, the onus of responsibility should be on the organization that provided the loan, tax return, disaster relief, etc., to confirm that it was, in fact, requested by the individual, rather than for the individual to somehow prove that they did not make the request.

Similarly, the relational web of responsibility means that firms that are third parties to fraud need to ensure that the other organizations and individuals they do business with are behaving appropriately. Consider imposter scams, where a fraudster uses a combination of your personal details and what they know about the government to impersonate a trusted source, catching you off guard.[13] Typically, these fraudsters have purchased lists of potential targets from data brokers and use publicly accessible communication and financial services to conduct the fraud. Firms that are third parties to the fraud have a responsibility to ensure that the products and services they profit from are used responsibly—and not for fraud.

When we shift our thinking about data responsibility from focusing on the data to focusing on the relationships between actors, we embrace the new narrative of the Indigenous metaphor. If we stretch the analogy of the Whanganui River, we might consider data to be the water flowing in our

digital landscape. The rivers this data flows through and the pools that it collects in are, like the Whanganui, actors shaping the landscape around them.

In this narrative, organizations that choose to interact with these data rivers and pools, seeking to source, transmit, aggregate, or consume data, should be considered custodians of the data, rather than owners. Their job is not to create multiple dams in the river and declare ownership of their stored water but to support the river to achieve its purpose. An organization collecting or transmitting personal data has a responsibility to ensure that the data is accurate. It should be possible, for example, for an individual to indicate in their credit history that they have no intention of applying for a loan. Or to remove themselves entirely.

A firm relying on personal data to provide a loan is responsible for ensuring that the application is valid, and they are providing the loan to the person they think they are. Similarly, a firm using bias-prone analytic solutions is responsible for ensuring that the analytic model is accurate and matches the intent and community norms of those if affects. Firms are also responsible for ensuring that their partners, suppliers, and clients, are meeting their responsibilities.

Moreover, a firm's permission to operate, its right to act, its agency, should be contingent on it upholding its responsibilities. To fail to meet responsibilities is a punishable crime in Indigenous Australian culture.

Compliance with regulation is often seen as a cost, a burden that must be borne, something to be minimized as much as possible. Data becomes a source of value when framed as property. It's the new oil—enabling firms to find customers, design better products, and target offerings more accurately. There's a chance that we have this the wrong way around.

Our responsibilities should be the source of value. Meeting one's responsibilities means that a firm has the opportunity to act, to find and capitalize on opportunities. If we all meet our responsibilities, we create a safe and productive environment for business. Hoarding or misusing data, ignoring one's responsibilities, should be a liability.

Just like the Whanganui River, our digital marketplace will quickly become polluted and unproductive if we focus on the individual rights of water holders, or our own selfish needs. An unpolluted digital marketplace is one where it's easy to do business—one where the data is pure, where we are aligned in our responsibilities to the clarity and flow of the river and its purpose, so that the river provides opportunity for every participant.

+ Endnotes

1. We are using "country" in the expansive meaning common in Indigenous communities, where it refers to the landscape, the actors that inhabit the landscape, and the network of relationships that bind them together. "Country is everything. It's family, it's life, it's connection." See: Australian Institute of Aboriginal and Torres Strait Islander Studies, "*Welcome to country*," accessed March 9, 2023. *View in Article*

2. Kate Evans, "*The New Zealand river that became a legal person*," *BBC Travel*, March 21, 2020. *View in Article*

3. It's entirely possible, and in fact, usually happens, that different parties will answer the question of "Who owns the data?" in exactly opposite ways. *View in Article*

4. Like boundary objects, data is pluralistic, with interpretation changing with context. *View in Article*

5. Most famous might be the story of a retailer predicting a teen's pregnancy. As with all these stories, details complicate the narrative. Such predictions are possible, but not as