

Identity is a phenomenon, not a property

A topic paper by Joe Andrieu (joe@joeandrieu.com) for the ID2020 Design Workshop, May 2016.

The use of the term “identity” and “identities” as selectable, definable, assertable properties undermines the ability for both laypeople and experts to communicate clearly about identity.

We all do it. We use the vernacular shorthand for a term because we know people will understand it. “Identity” is one of the 3000 core words in the English language. It’s powerful and resonant. Essentially all English speaking persons understand it.

Yet when we speak about identity as a property of a person or thing, it is easy to gloss over the mechanics of what’s going on and obscure the very issue we are trying to discuss. As a result, different people use the same term to mean radically different things and smart, well-meaning collaborators get lost in miscommunication.

Sometimes, “identity” means a specific human individual. In others, it means what someone wants to be called. Sometimes it **must** be issued by an authority, sometimes it is self-asserted. Sometimes, “identity” refers to the cultural association with a particular class or group, e.g., my identity as a straight white male. These are all valid. When used in technical conversations, this innate flexibility makes it hard to discuss identity systems and how proposed solutions may or may not address particular needs.

This vernacular use of “identity” is a lot like using “inertia” when we really mean “momentum” or when we talk about the “gravity” of an object when we mean its weight.

The physicists among us understand that “inertia” doesn’t increase or decrease, “momentum” does (if you accelerate a mass). Both laypeople and experts can use the terms interchangeably and, usually, nobody gets hurt. Such use lacks rigor, but it rarely confuses.

Similarly, we say that the Earth has a “higher gravity” than Mars, when what we mean is that the force of gravity from the Earth on an object is higher (on the surface of the Earth), than the force of gravity from Mars would be on that same object on the surface of Mars.

In most cases, the shorthand terms allow us to speak quickly without worrying about the details, so that the conversation can focus on what really matters—how cool it would be to walk on Mars!

But when you start engineering a spaceship that must launch on Earth and land on Mars, it pays to be precise about the force of gravity and about how much momentum that ship will have when it decelerates against Martian gravity.

Similarly, if we are building, defining, or fixing identity systems, it pays to be rigorous in our language.

Identity is correlation.

Gravity is like Identity in another way. Both exist only in relationship. Gravity is a force exerted by two objects on each other¹. An object doesn’t have gravity except in relation to another object; (Newtonian) gravity is not an innate property of an object. It is an emergent phenomenon between two objects. In

¹ Or in the case of objects with multiple particles (e.g., EVERYTHING), a force an object exerts on itself. But such gravity is irrelevant when discussing the effect of gravity elsewhere.

the same way, identity cannot exist on its own.² Identity only occurs when one party correlates another party with a label or previous interaction. Let me repeat that.

Identity occurs when one party correlates another party with a label or previous interaction.

We can identify an individual by their name or citizenship or by the fact that you've seen them before.

The eye witness who stands up, points to the accused and says "That is the person I saw kill the victim." That's identity. When the witness correlates the physical person in the room with the person with whom they had a prior interaction, that's identity. No identifiers required.

Friends regularly remember other friends' romantic dalliances by keeping track of the memorable bits from the stories they hear. That's identity. When a person gets stuck with being known as "the pink-eye guy" because he was the one who had pink eye on the first date with your friend, that's identity. In situations like this, the identifier—and identity—is *inherently* out of the subject's control.

This correlation between the subject and something else... THAT is the core of identity.

Using the term identity without clarity around what is correlated, by whom, consistently confuses people in just about every identity conversation I've been in.

When building systems we necessarily use identifiers (see Joe's³ and Randy's⁴ posts) in various ways to correlate a current user with others (including their past use). Often those identifiers, get referred to as "identities". We refer to our username or handle as an identity. We say users select "an identity" when using the Information Card card selector.

We know that the card itself isn't the identity,⁵ but our vernacular often treats it as if it were.

My goal is to convince you to stop using such shorthand when discussing identity systems in both technical conversations and when describing identity to laypeople. If you use the vernacular when touching on identity in other conversations, fine. But when the point is to clearly and concisely discuss identity, don't do it. There's a better way.

Identity management is correlation management

When we speak of government "identity", we mean the facts that a government authority **correlates** with a particular individual. When we speak of corporate identity, we sometimes mean the trade dress and associations that customers and competitors **correlate** with a given corporation. When we speak of a self-asserted identity, we mean facts that an individual wants to have **correlated** with them by others.⁶

² Again, we can construct an "internal" identity that only the possessor knows about, e.g., my secret name, but as long as it is secret, it has no bearing on anyone else.

³ Andrieu, Joe. "The Identity Quartet" *My personal space*. Blog. March 3, 2009. Online. Accessed 2016.05.19 <http://blog.joeandrieu.com/2009/03/12/the-identity-quartet/>

⁴ Farmer, Randy. "The Tripartite Identity Pattern" *Habitat Chronicles*. Blog. October 17, 2008. Online. Accessed 2016.05.19 <http://habitchronicles.com/2008/10/the-tripartite-identity-pattern/>

⁵ We should realize that. But often even experts confusingly refer to the artifacts of identity as "identities"

⁶ Whether or not they choose to use that means of correlation is a different matter.

Instead of referring to identity as if it's a property, speak about the specific actions individual take (or data inferred, sensed, or sampled) and how the observing party correlates that with something else.

For example "an Individual selects an information card, and the relying party correlates that card's claims with the current user account."

Or "The user selects a profile; the software uses that profile's key to sign the challenge; the relying party validates the signature, and correlates the blockchain profile with the current user."

The ambiguous term "identity" wasn't required, nor did we refer to "an identity" as if it were a property.

Understanding correlation is understanding identity

When building identity systems, the crux of both design and implementation is providing the needed correlations *without* enabling undesirable ones.

The idea of putting all newborns' footprints on "the blockchain" makes sense for public policy staff trying to map their world to the latest fancy, all-promising technology. Those of us more familiar with the public nature of blockchains understand such an approach would expose too much personally identifiable information. My contention is that the easiest way to correct these understandable mistakes is to stop treating identity as a property and to discuss correlation instead.

The key to understanding the requirements of an identity system is to understand what needs to be correlated and what correlations should be impossible or economically infeasible. These use cases and non-use cases, when couched purely in terms of correlations, can be used to design and evaluate potential solutions. Further, if we don't have a reasonably complete set of correlations and anti-correlations, it's likely that any proposed solution with either fail to do the job well or cause more problems than it solves.

To correlate is to identify

Identity *is* correlation.

It's that simple.

Rather than treating "identity" as if it were a property that exists independent of the subject and observer, discuss how to correlate a given individual across contexts in ways that enable access to services while preserving human dignity and freedom.

Do that, and we'll all have a much easier time co-creating a viable solution to digital identity for every last child on Earth.