



[Tech Notes are short articles discussing library-related technology.]

[Tech Note:] Identity Crisis

Philosophy has something to say about everything. One of my philosophy professors penned a paper that purported to show that a fetus and a prisoner on death row had differing moral claims to life based on their *personal identities*. I'm not sure how good the paper was, but it certainly grounded philosophy in the real world.

Identity has been understood to be what I say about myself and what others say about me.[1] Perversely enough, what I say about myself ("I'm the King of England") is less trusted than what others say about me ("He's this guy from Wisconsin"). Identity in this sense is equivalent to reputation.

A more prosaic but important problem is identity in the cyberworld, particularly on the web. Mostly the way we have identified ourselves to applications and services has been through a username and password. This is a means of authentication—proving we are who we say we are. Beyond authentication is authorization, being allowed to access or use a set of resources given who we are or what group we belong to (e.g., professor, student, library patron, library staff person). Many sites or databases limit or disallow resource access without prior authorization. Some have called this Identity 1.0.

Identity 1.0 poses problems: it is inherently insecure given that people may, in many cases, provide passwords that are laughably easy to crack. Even where there is an attempt to make people provide stronger passwords the user is still a weak link: the more difficult to remember passwords are, the more likely they will be on a piece of paper stashed somewhere nearby.

The authorization piece also can present problems as more and more categories of access open up. As noted above, there are many roles that people assume when they log in to a system. Access to resources is not always as fine-grained as we would like, e.g., determining which specific databases what are open to a particular person.

Portability is also a problem. Suppose I build up a list of books and a history of purchases on Amazon. How can these go where I go, e.g., LibraryThing or, better yet, my local library? Right now I can't.

Still another problem is privacy. Sites almost always guarantee privacy, but come up lacking when their feet are put to the fire. There is no physical inhibition against releasing personal information.

In spite of these limitations, Identity 1.0 is still the dominant method of identification these days. But what comes next?

Think of the way that your driver's license or state ID card works [2]. Walk into a liquor store. Gather up your intended purchases. If you look too young to the clerk, or the clerk is particularly punctilious, you may be asked to produce identification. You hand her your driver's license. She examines it. She is looking for an indication you are who you say you are (your photo), that a legitimate authority has issued it (e.g., the State of Wisconsin), that it is genuine, and that your date of birth indicates you are old enough to purchase liquor. When she's satisfied she sells to you. Note that an independent, disinterested, trusted entity has provided the credentials to identify who you are and certify your age. Notice that there is a separation between the acquisition of the ID and its presentation. It can be used for various transactions any time after it's issued. This makes for potentially extremely large scale (think of all of those liquor store and other transactions) and privacy (the issuer does not know when or how often I use it).

Note a few interesting aspects of this transaction: the clerk does not contact some state entity to perform the verification; the consumer carries one card that works at any store; there is no relationship between the issuer and the store. Wouldn't it be nice if the myriad vendors and information sources on the Internet offered similar capabilities from a single credential? Seamless authentication and authorization appear to be goals for all of the 2.0 projects.

Microsoft Passport was an early entry into this area, well before the Identity 2.0 moniker. Microsoft envisioned it as a method for consumers to have a single sign-on to multiple merchant web sites, making their purchases seamlessly. It never gained enough momentum in the marketplace to become a success. Now Microsoft is back with Windows Live ID, which according to a Wikipedia article has its own problems [3]. The idea is still single sign-on across a wide range of sites. But of course Microsoft has little interest in the library world—it's about the money. And for whatever complex of reasons, there is a degree of antipathy for Microsoft among potential sites for adoption of Live ID, making widespread adoption of it problematic except for Microsoft-owned sites.

A more interesting scenario from a library aspect is a single sign-on to all of the resources of a college campus, including the library. This might be implemented by a "smart card" that carries a balance of funds for use at the cafeteria or bookstore, plus acting as ID for access to campus buildings, and to online campus services including the library. This is analogous to driver's license/state ID, but with a richer set of capabilities. Ideally, of course, we would like to have students/faculty/staff be able to use the same card at the local library, and conversely for local residents to be able to use their library card for access to library resources at a campus. This, alas, is something for the future, but is actually realizable with current technology, although it would undoubtedly require a costly investment and lots of agreement among participants to make it happen on a large scale.

The use of a single identity authentication across multiple service providers is called *federated identity*. Figure 1 shows the relationship among the parties involved. As noted on the diagram, trust must exist between the identity provider (IdP) and the service

provider (SP). This is analogous to the trust exhibited by liquor store clerks in the State of Wisconsin. While this is superior to a username/password sign-on for each resource, what would be even more desirable is authentication without implied trust between the IdPs and SPs. There is a project underway which aims to provide a piece of the sort of relationship shown in Figure 2. Here I claim my identity by providing an identity URL to the SP, which in turn sends it on to an IdP. In turn, the IdP sends back to the SP a confirmation of my identity.

Notice that in Figure 2 there is no trust relationship between the IdP and the SP. Rather than relying on the Identity Provider being a trusted entity, the Service Provider relies on the authenticity of the identity URL. It must be of a form and format that will yield a valid response from the Identity Provider.

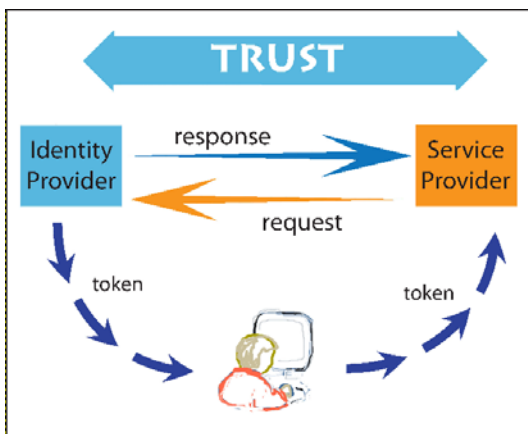


Figure 1.

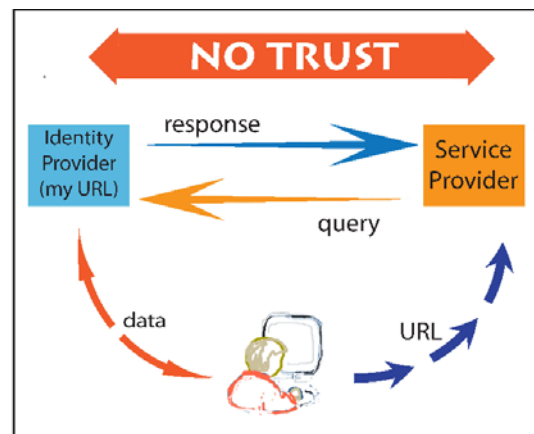


Figure 2.

Shibboleth, widely adopted by higher education institutions and to be included as a piece of Internet2, is a federated solution. Shibboleth “is a standards based, open source software package for web single sign-on across or within organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner.”[3] Shibboleth has not found wide adoption outside of the higher education community, so it is unlikely to represent a universal solution to the single sign-on problem.

The authentication landscape is in quite a bit of flux right now, with more companies and open source organizations trying to provide solutions to identity that require less trust between the parties and more power to the user. So it may not be the best time for an enterprise to commit to a single vendor and technology.

Takeaways From This Article

- Authentication and authorization at most sites are in a primitive state.
- The driver’s license analogy may be useful to think about good authentication—third party provides authentication, authenticator need not trust user.
- Single sign-on is the ultimate goal of Identity 2.0.

- Single sign-on requires co-operation (e.g., adoption of standards) among all parties.
- Identity technology and protocols are constantly evolving.

[1] This Tech Note owes much to two presentations by Dick Hardt, the first at the O'Reilly Open Source Convention in 2005 (find it [here](#)) and the second at the Emerging Technologies Conference in 2006 (find it [here](#)). (Links tested 6/25/08)

[2] *Ibid.*

[3] Shibboleth at <http://shibboleth.internet2.edu/> . (Link tested 6/25/08)

—Tom Zillner (tzillner@wils.wisc.edu)