Ralf Bendrath

Digital Identity—Bug or Feature of Web 2.0?

"The Internet was built without a way to know who and what you are connecting to". This sentence opens the manifesto "Laws of Identity", published by Microsoft's Chief Identity Architect Kim Cameron in May 2005. This is actually nothing new, as online anonymity had been a hot topic of debates since the early Usenet. In 1993, Peter Steiner became famous among netizens with his New Yorker cartoon "on the Internet, nobody knows you're a dog". If we agree, the obvious next question will be: Is this a bug or a feature? While the dogs of the 1990s liked this feature, Kim Cameron and many others see it as a bug.

What is identity in the first place? A minimal definition would include the fact that one thing or one person is the same after a bit of time has passed. While this sounds easy, philosophers from ancient times to today have had a hard time to think it through. Buddha already asked:

„Years ago you were a small baby; then, you were a boy; then a youth, and now, you are a man. Is there any identity of the baby and the man?"

Too complicated for today's IT world? Let me introduce you to Alice, Bob and Eve, who have become famous because some cryptographers wrote about their stories. Alice and Bob had an affair at high-school ten years ago, but then Bob moved to Linz, and they lost contact. Bob is now in love with Eve. One day, Alice comes to Linz for an electronic arts conference, goes out in the evening, and sees a guy at the bar. She thinks: Is this my lost high-school-boyfriend Bob? Now, if he really turns out to be Bob, is he still the old Bob she used to go to the movies with? What if they still like each other after all these years, and Alice ends up drinking with Bob? And what if something happens that night that Bob might regret the next morning, because Eve, his girlfriend, should not find out about it?

British philosopher John Locke already thought about this in 1690. In his "Essay Concerning Human Understanding", he asked: "Is not a man drunk and sober the same person? Why else is he punished for the fact he commits when drunk, though he be never afterwards conscious of it?" Taking this to the online world, we can ask: Who are you when you go online? Are you still the same person as in the meat space? What if Alice goes to *Second Life* and meets Bob there?

Web 1.0 and Web 2.0—from linking documents to linking persons and contexts

In Web 1.0, we did not really ask these questions. It was considered a great document server, an enhanced version of FTP. Matching your real identity to what you did online was not an issue. Instead, anonymity and pseudonymity were important, and they were cool and opened new options for experimentation. For example, Bob could play Alice's role online and figure out how it is to be a woman. And Eve could pretend to be a man—or even a dog. Or was Eve played by a dog?

Of course, there are two underlying identification systems built into the Internet. The most fundamental one is the IP address space, which allows computers to identify each other and ensure their connections. In the time before dynamically assigned IP numbers and Internet cafes, they could basically be used as an identification system for people. You only had to find out which computer had which IP number (which of course is not too easy if the administrator of a certain IP range is at the other end of the world and not willing to cooperate). The other identification layer is the Domain Name System, which allows the unique naming of email addresses, web sites and other services. Here, the ongoing and long-lasting battle within ICANN over who can access the WHOIS database for which purposes shows us that even in Web 1.0, anonymity has never been total and is heavily contested. But as a rule of thumb, with dynamic IP addresses and anonymous or proxy-registered web services, you can not be sure that you find out whom you are connecting to.

With Web 2.0, this has changed. It is not about linking documents anymore, but about linking persons. People seem to have a desire to match their online and offline identities, to link them with others in a certified way, and to tell the world about it. And people seem to want to link—humans are social animals, after all. This is probably one of the reasons why these networking platforms, be they *MySpace, Xing, Friendster, Facebook,* or even the notorious German *StudiVZ,* are so popular.

What you build in these platforms is not just a social network (or more correct: a technical map of a social network), but also a reputation and an image. What does your picture look like? How many friends do you have? What do they write in your guestbook? Is your layout cool enough? You also build a reputation as an Amazon reviewer or an *eBay* seller, as a *Flickr* photographer, and so on. So, one day Alice thought: Wouldn't it be cool to take my high *Slashdot Karma* and use it for *eBay?* Or use my *eBay* reputation for *MySpace?* She figured out that there is no real standard for interchanging reputation. In the end, it is because Alice's reputation on *Slashdot* does not say much about her reliability as an *eBay* seller. In fact, the way these identification systems are built only allows their users to establish their identity and social relations within their systems, effectively turning them into monolithic silos. Again: Is this a bug or a feature?

Sociologists call this "functional differentiation", and it's the foundation of complex societies. Georg Simmel in his 1907 essay "The Secret and the Secret Society" already described in detail how important it is that different people know different things about you. Bob's girlfriend Eve knows other things about him than his boss, and his banker knows other things than the members of his bowling team. You play different roles in relation to different people. And sometimes it is even important that these roles are not linked to each other. Similar to Simmel, Helen Nissenbaum has recently tried to build a new normative foundation for "privacy as contextual integrity": Information about persons is generated everywhere, but privacy is the fact that these do not spill out of their context—or at least that this transfer is controlled by the individual affected.

But what if people want to link the different contexts they are active in? There is no defined interface for linking your personal profile at e.g. *Xing* with your *MySpace* account. You can use your "Google Checkout" account for logging into several Google-owned services, but you cannot use it for Yahoo or other systems. This is why there is a growing trend towards open identification standards and protocols for the Internet that give the user more control and free him or her from the limits of these walled corporate gardens. The discussion around this is being held under the label "user-centric identity".

Digital Identity 2.0—technology, emerging protocols, and privacy

User-centric identity approaches include low-tech HTML tweaks (microformats) like "vCard", a machine-readable business-card people can put on their website that contains meta-data tags about themselves, their contact information, or their institutional affiliation. These are of course easy to forge. Anybody could put a vCard on his website and pretend to be me. If the information in the vCard is correct, on the other hand, this is mainly a great help for spam-bots. Moreover, you can not tell if a vCard for a specific John Doe is referring to the same person as a different vCard at a different website that includes different contact information for a John Doe. It might be the same John, but with a different company, or it might be a totally different person. What is missing is a defined address space here, just like the domain name system. More sophisticated approaches therefore use methods similar to URLs as identifiers. "i-names" is such a project, based on Extensible Resource Identifiers (XRI) developed by the OASIS group. You can register your i-name with a network of i-brokers, like you register your domain name with DNS registrars. These third parties are generally called "identity providers". They confirm to other services that I am the correct person to use a specific name or identifier, or simpler: They identify me towards other parties. In exchange, of course, I have to rely on these trusted third parties for each transaction I do.

There are quite a few similar protocols and approaches out there now that have developed over the last few years: MicroID, Lightweight Identity (LID), OpenID, Yadis, Secure Authentication Markup Language (SAML), ID-WSF, Windows Cardspace, Higgins, Shibboleth, and more. Some of these, like OpenID, MicroID, LID, and Higgins, have emerged out of the blogger and Web 2.0 community. Others are driven by large IT corporations like Microsoft (CardSpace) or the Liberty Alliance that includes Sun, Oracle and others (ID-WSF). A few also come from one of the several formal or informal standards groups for the Web, like OASIS (SAML) or the ITU (X.509). Because of all these different approaches, some standards organizations, like the W3C, ISO and ANSI, have set up working groups on identity management recently, so we can expect some interesting standards wars here in the near future. Currently, the most interesting approaches in the Web 2.0 context are Microsoft's *CardSpace,* which is being shipped with Windows Vista, and *OpenID,* which has emerged out of the loose network organized around the "Identity Gang" and the Internet Identity Workshops. Here, the technology design is extremely important, because it can make a huge difference in terms of traceability and linkability. It is interesting that Microsoft's approach "CardSpace" is in fact much more privacy-friendly than the community-driven standard OpenID. While OpenID is like a light-weight single-sign-on service with the ID provider sitting in the middle and knowing all transactions, in the CardSpace model the ID provider does not necessarily have to know with which service I am connecting, and the different roles I play in different web services can be strictly separated with different and un-linked identifiers.

Name laws and digital ID cards—government as the ultimate identity provider

At the center of the debate around digital identity management are the identity providers. They register my unique ID, they identify me towards others, and they may be able to track all my activities. Corporate identity management systems have done identification of their employees for quite a while. They use role modeling, or provisioning, for differentiating the several tasks their employees can take. Who can enter the premises? Who has access to which database? Who can authorize buy and sell orders for which amount of money? This is where the big players like Oracle, Novell, or Sun, come from. With the end of the closed firm and the emergence of web-based collaboration, corporate identification systems are merging with web-based ID standards. The use cases of identifiers here are normally called "provisioning", "identity federation", or "workflow auditing". In the end, of course, it is about controlling employees. And this is one of the most fundamental functions of identity management: Control.

But there is one much more important type of actor, whose role in the identity space is often overlooked and ignored. Who was the first to establish identity management and identification systems? It was the early modern state. In the 15th century, the first laws were enacted in Europe that made it illegal to change your name during your life. Now, for making sure to others that you are the person whose name you pretend to have, you need some extra proof in the form of identity tokens. First, these used to be official letters or seals, and in the last century we saw the development and spread of passports and ID cards. Nowadays, governments try to link the passport (which certifies a name) more closely to the physical body of each citizen by using biometric technology. But the need for a more fixed coupling of the identity tokens (identifiers) and the persons (those identified) was again seen much earlier. Jeremy Bentham, who invented the idea of the Panopticon, in his essay "The Principles of Penal Law" 200 years ago, suggested that every citizen should have his name tattooed on his arm. This inscription of an identifier in the body of the identified is merely a more radical version of putting fingerprints into passports. In Bentham's time, biometric fingerprint readers or iris scanners were not yet available, so a human-readable identifier was the natural solution.

But even today, your tattoo is not transmitted when you go online. So, some governments now want to establish a certified, official link between your real physical identity and your online identity. The government agencies that issue passports and ID cards would then become identity providers for your online life, too. The difference is: The government does not know whom I show my passport. But with most existing digital ID management systems, it would necessarily be part of the transaction when I identify myself towards a third party with a government-issued digital ID token. This obviously holds quite some potential for large-scale surveillance and control of online behavior. Especially in countries without a tradition of ID cards like the U.S. and the U.K., people resist the idea of mandatory online identification.

So, what do you do as a security politician when you want to set up such a system? You start with groups like foreigners or criminals. The U.S. Senate has a bill pending right now which would force all convicted sexual offenders to register all their email accounts and all other online identities with the authorities. They are dead serious about this: If people fail to register, they will face up to ten years of imprisonment. This is not for raping anyone; this is just for not telling the government all their online user names and pseudonyms.

Other countries with more of a surveillance and control tradition have started building up online identification systems that would force Internet users to register with their real name before using these systems. South Korea is currently developing an "Internet real-name system" for bloggers that they would be forced to use for posting blog entries and commenting; the People's Republic of China is working on a "real name verification system" for bloggers, but also for online games. This again shows the control function of ID systems. Besides tracking and profiling people, it also allows for zoning the net and providing automated entrance-control. Similar plans are under way in *Second Life* for the virtual red-light district.

Germany is currently working on its reputation as the land of more advanced bureaucracy. The „E-Government 2.0" program", published by the German Interior Ministry in September 2006, has an interesting chapter on electronic ID cards and "e-Identity". The federal government plans to issue an electronic ID card from 2008 on, which will enable people to authenticate themselves online with their government-certified ID. So in Germany, registration of your online identity with the authorities is not a "for criminals only" thing. It will apply to the whole population.

In the end, we might end up with the government as the ultimate trusted third party, and get a perfect ID management system that encompasses everybody. Many e-commerce businesses would certainly like and ask for this digital ID card in order to prevent fraud and other unwanted activities. Depending on the technology design, the government as the digital ID provider then might be able to track people's behavior online. The government would become the ultimate trusted third party. Again: Is this a bug or a feature?