Facilitating online integrity using *OpenID*

Tony McDonald
Learning Technologies for Medical Sciences (formerly the Faculty of Medical Sciences Computing)
The Medical School, Newcastle University, UK

"Integrity is the essence of everything successful" (Richard Buckminster Fuller)

How can individuals behave in an online environment that enables them to keep and enhance their integrity and reputation, fills colleagues with confidence, allows for free exchange of information and yet still manage to keep private those things that they wish to keep private? The Internet is a largely anonymous world, with flame-wars, spamming, phishing and distrust being the order of the day, but I will argue that by allowing some carefully controlled non-anonymity to 'leak out', mutual trust can be built up. I will outline some existing authentication and authorisation systems, and will touch upon identity management issues in general. I will discuss some initiatives such as Shibboleth, OpenID in general and discuss in particular how OpenID might be used in practice to allow individuals to have a greater say in what information is held about them, and how this information is used by third parties.

Keywords: integrity, persona, authentication, authorisation, shibboleth, openid, identity management, privacy

Introduction

It is a truism that the Internet has made anonymity highly available, and there are many very good reasons for wanting to have the option of anonymity when it is required. However, one of the downsides to this anonymity is when you want people to take you seriously, or trust you, or listen to a point you're making because you happen to have some experience in that area. If they have no way of knowing who you are, this becomes exceedingly difficult.

I want to outline a way forward that allows educators, students and institutions to behave online in a manner that gives the potential for them to enhance their integrity, based on the requirement of some level of identification of users of systems. The difficulty is in ensuring that as much control of the personal and private information about a person or group is under that entities' control. The good news is that there are initiatives under way at the moment that leads me to conclude that this goal is not insurmountable.

Why do I have to be identified anyhow?

Put simply, trust comes from knowing who you are talking to - people rarely unconditionally trust someone they hardly know, and people behave better in environments where they are known. Trust leads to sharing of knowledge, and in doing so, all participants' benefit. Therefore, should you wish to gain someone's trust online, then some aspects of your persona will need to be made available for people to peruse, but the crucial point is that you should be able to decide just what should be shared about you. Most systems today do not make that distinction, preferring to control all information held about you, or a group.

The traditional method of identifying yourself with an online system takes place in three steps;

1. You authenticate yourself with a system (usually with a username/password, but keycards and other physical devices can be used, either singly or in combination with a username/password pair) - *authentication*
2. The system checks your credentials and checks what you are able to do within the system - *authorisation*

3. You use the system, and depending on the roles designated to you by the system, you are able to do or not do certain tasks within the system. Data may be shown to you or hidden from you, and you will generally interact with the system in many different ways.

Dependent on the type of system being used, you may find you are able to access similar systems without needing to authenticate again, or you may have authenticated at a *portal* system which merely acts as a gateway to other authenticated services. This is called 'single-sign-on'[1] or SSO.

In the way in which we are using the words *authenticate* and *authorise*, it is important to note the distinction between the two.

1. Authentication[2] is "the act of establishing or confirming something (or someone) as authentic, that is, that claims made by or about the thing are true". From an access to information standpoint, this could be as simple as the ubiquitous username/password combination. Note that having a correct username and password combination does not imply that the username used is the one that was allocated to the person you think it was. You can think of authentication as having the key to a car.
2. Authorisation[3] is described as "the concept of allowing access to resources only to those permitted to use them". You can have a perfectly valid username and password to a system, which would allow you to post messages to an online forum and edit details of events, but may not be allowed access to areas holding personal information. In the car example, authorisation is where you are allowed (probably by law) to drive the car.

These two terms are not generally interchangeable, although it is common to see 'authentication' being used to imply authorisation as well.

With these definitions, we are ready to look in more detail at how information about a person or group is managed by authentication and authorisation systems, and how some of the more common systems work. This is generally known as Identity Management.

Identity management

Wikipedia defines Identity Management[4] as "Computer scientists now associate the phrase, quite restrictively, with the management of user credentials and the means by which users might log on to an online system."

One question arises from this definition, *Who says I am who I say I am?* It is one thing to be given a username and password, but there is no guarantee that matches the legitimate owner. It therefore becomes clear that there needs to be a joint agreement on who the authority is who is stating that you are who you say you are. These authorities and systems are generally known as Identity Providers[5]

Some typical Identity Providers are;

1. Host University, Professional Society (e.g. Royal College of Surgeons, GMC, AMC)
2. Bank/Building Society/Mortgage holder
3. 'Trusted' corporation/company (e.g. Google, Microsoft)
4. Government

Identity Provision systems come in various guises; they can be centralised, as you might expect to see at a university. These systems have been generally designed to be of use to the central administration of the institution rather than the users of the systems. This means it can be quite difficult to change personal details, and in addition, the individual users of the system may have little say in what attributes (such as email address) are held about them, and to what uses it is made.

A centralised system is generally not able to share authentication details with other systems, and is liable therefore to act as a data island. This may be perfectly acceptable to the user if the size of the 'island' is large enough (Google[6] uses the same login for web search services, email, news groups, calendar information, document storage, blogging services, map information and a host of other resources, but even Google is not the whole Internet). Expanding the size of the authentication base means more people can use it to authenticate, but increases the vulnerability of the system to outages, deliberate attacks and

phishing[7] attempts and security breaches become of greater importance (an example in England in late 2007 - 25 million personal records lost)[8].

In the spectrum of centralised-decentralised authentication systems, the middle ground is held by loosely coupled authentication systems, such as Shibboleth[9], which uses the concept of a 'federation' to allow users to authenticate at one site, and be automatically authenticated at subsequent sites within the federation (this is an example of 'single-sign-on'). There is some leeway in the way in which users may be authenticated at any one site, but attributes, such as email addresses are passed between systems in proscribed ways. In theory it is possible for individual users to limit the attributes that are passed around between these systems, but in practice this does not happen often, and it is likely that institutional procedures will be adhered to when it comes to the transfer of attributes between authentication systems.

Generally speaking all of these systems allow little freedom for the user to define what attributes are passed around between them, and this is primarily because the end-user has little or no control over the Identity Provision system.

One system that does allow this flexibility is OpenID[10]. OpenID allows an individual to use a single digital identity across OpenID-enabled services. The source code is Open Source (as is Shibboleth) and there are many OpenID providers that you can choose from to host your identity information (compare this with the Shibboleth and central authentication systems, where you have no little or no choice as to where your identity information is held). OpenID uses profiles to allow fine-grained access to your attributes, and is granular enough for the individual to be able to allow, say, their surname to be made available to one online forum, whilst allowing first name, surname, sex and email address to be made available to another. What this means is that a user can decide what information is shared between online services, and it is up to those services to either; (a) provide the service as best they can - even if a specific piece of information, say email address is *not* permitted to be shared or (b) tell the user that with the best will in the world they can't use the system without that information. The onus is then on the user to share that information or not, but all parties are aware of their responsibilities in the transaction. OpenID is starting to get a large amount of momentum behind it with Google, Microsoft, IBM, Yahoo and VeriSign all sitting on the Corporate Board of the OpenID Foundation[11]. MySpace have recently announced their intention to make some of their services OpenID compatible, increasing the potential number of OpenID-enabled accounts by over 500 million (AOL, Yahoo and others have already said they will support OpenID) [12]. Both Microsoft and Google have announced[13] their intention to support OpenID 2.0 on their platforms, with Microsoft enabling all 460 million users on the LiveID platform and Google are making all google accounts OpenID compatible.

An OpenID is not a username/password combination - it is a URI, for example http://tonemcd.com/ is one OpenID that I use and I can use that to access services that are OpenID enabled. Because an OpenID is a URI I don't have to worry about creating yet another username at a site and try and remember that. I am able to use the URI http://tonemcd.com because of a feature OpenID has called *delegation* where, I have put code in the HTML of my webpage at tonemcd.com;

```
<!-- openid delegation -->
<link rel="openid.server" href="http://www.myopenid.com/server/" />
<link rel="openid.delegate" href="http://tonymcdonald.myopenid.com/" />
```

This HTML will redirect any requests for OpenID authentication to my current OpenID provider, www.openid.com. I am able to do this because I own the domain tonemcd.com.

This opens up the possibility for universities to set themselves up as OpenID providers which would leave them completely free to assign every member of staff, every student and every arbitrary group in their group their own OpenID (eg for Newcastle it could be namm2.ncl.ac.uk, or nanogroup.ncl.ac.uk). MIT[14] works in a slightly different way, http://auth.mit.edu/your_username_here, whilst Brigham Young University[15] operates slightly differently. Because individuals and organisations would trust that university, they would know that if I were to use that particular OpenID online (namm2.ncl.ac.uk), then I would be authenticated as being from Newcastle University. I could then set up a multitude of profiles at Newcastle, ranging from the very lightest (first name or nickname) through to all my demographic data, etc. The onus is on me to use the correct profile in the appropriate place, and it is my reputation that is in

jeopardy should I choose poorly; "It takes many good deeds to build a good reputation, and only one bad one to lose it" - Benjamin Franklin.

The previous example could be extended to other organisations and professional bodies such as the General Medical Council in the UK (GMC[16]), who could use it to authenticate users as having an affiliation with that body, as they would carry out the authentication/authorisation based upon details they have available about the user (this may include signed documents and certificates). This then could be used as the basis of an online community of practice in that discipline.

OpenID in action

Using OpenID in action is very simple. Instead of entering a username/password combination, I use an OpenID;

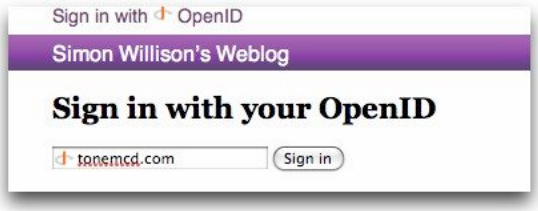

Figure 1: Authenticating with OpenID at a site

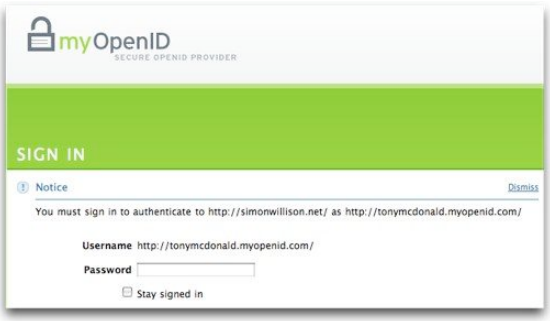

Figure 2: Delegation happens, I am sent to myopenid.com (my current OpenID provider) and provide a password

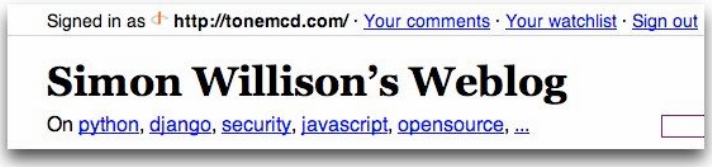

Figure 3: I am redirected back to the original site, and can now login to other OpenID-enabled sites.

Benefits/pitfalls and privacy

Trusted organisations can become OpenID providers and people/groups associated with them can then be associated with that organisation. This provides a certain level of integrity as a matter of course. People are free to use whatever profile or persona they like in their online dealings, but it is likely that if you were using a university OpenID, certain professional standards would be expected.

OpenID is susceptible to phishing attacks, as indeed are many sites such as eBay, PayPal, Google etc. but one of the main issues with OpenID is that it works quite differently from sites that people are used to visiting.

OpenID allows you to specify profiles that can be as detailed or as sparse as you like, allowing you to be as private as you like. The downside is if the service you are trying to use requires some personal

information you may find the experience of using the site is marred if these attributes aren't shared. However, the onus is on you to provide the information - it is not provided as a matter of course.

Discussion and conclusions

Integrity and at least some level of authentication go hand in hand. The balancing act is in ensuring that, as far as possible, the data owner is in control of their own data, and that the service the person is accessing has enough information to be able to provide a worthwhile experience to the end-user or group. There are systems available that allow access to this information, ranging from the highly centralised to the highly decentralised, and they can be viewed as alternatives to OpenID, but I would argue that few have the traction that OpenID is now gathering, and in any case can be integrated into OpenID authentication/authorisation schemes[17], which helps immensely in integrating OpenID with existing Learning Management Systems.

Decentralised systems such as OpenID afford the end-user the luxury of being able to restrict the data that authentication systems ask for, and to do it in a highly flexible and open manner. It is more of a pact between the user and the service provider than other, more centralised systems.

By having systems that handle user information on the users terms, but which can be backed up by the integrity of an institution such as a university, professional body, bank or large corporation, it is likely that the online integrity of educators and students will increase.

References

1 http://en.wikipedia.org/wiki/Single_sign_on [accessed 29 July 2008]

2 http://en.wikipedia.org/wiki/Authentication [accessed 29 July 2008]]

3 http://en.wikipedia.org/wiki/Authorisation [accessed 29 July 2008]

4 http://en.wikipedia.org/wiki/Identity_management [accessed 29 July 2008]

5 http://asc.gsa.gov/portal/template/faq08.vm;jsessionid=9E7A383264A5CE4D091F5019BEDAB157 [accessed 29 July 2008]

6 https://www.google.com/accounts/ManageAccount [accessed 29 July 2008]

7 http://en.wikipedia.org/wiki/Phishing [accessed 29 July 2008]

8 http://www.guardian.co.uk/politics/2007/nov/21/immigrationpolicy.economy3 [accessed 29 July 2008]

9 http://shibboleth.internet2.edu/ [accessed 29 July 2008]

10 http://openid.net/ [accessed 29 July 2008], see also http://www.ariadne.ac.uk/issue51/powell-recordon/ [accessed 3 Oct 2008]

11 http://openid.net/foundation/ [accessed 29 July 2008]

12 http://www.techcrunch.com/2008/07/21/myspace-to-join-openid-bringing-total-enabled-accounts-to-over-a-half-billion/ [accessed 29 July 2008]

13 http://openid.net/2008/10/30/microsoft-and-google-announce-openid-support/ [accessed 3 Oct 2008]

14 http://auth.mit.edu/ [accessed 3 Oct 2008]

15 http://wiki.eclab.byu.edu/index.cgi?OpenIDForBYUDevelopers [accessed 3 Oct 2008]

16 http://www.gmc-uk.org [accessed 3 Oct 2008]

17 http://blog.case.edu/jeremy.smith/2007/03/09/openid_server_integrated_with_cas [accessed 3 Oct 2008]

Author: Dr Tony McDonald, Assistant Director, Learning Technologies for Medical Sciences (LTMS), Faculty of Medical Sciences, The Medical School, 16-17 Framlington Place, Newcastle University, NE2 4HH, UK. He is responsible for IT provision across several degree programmes, primarily Medicine, Dentistry and BioMedical Sciences in the Faculty of Medical Sciences. This includes teaching and learning systems along with the administration systems to support them. He has worked on several identity, learning technology and ePortfolio projects, at the institutional and national. Email: tony.mcdonald@ncl.ac.uk

Please cite as: McDonald, T. (2008). Facilitating online integrity using *OpenID*. In *Hello! Where are you in the landscape of educational technology? Proceedings ascilite Melbourne 2008*. http://www.ascilite.org.au/conferences/melbourne08/procs/mcdonald-t.pdf