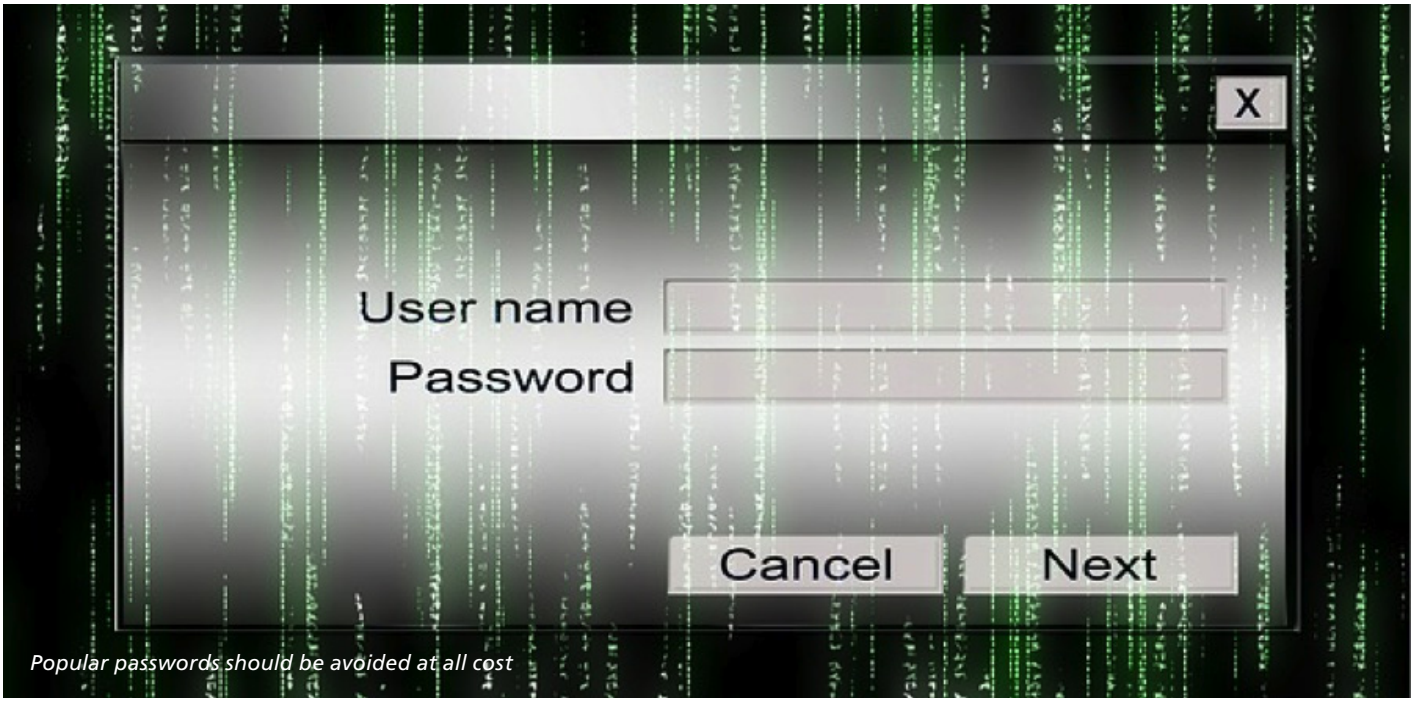
e science TALK
e-ScienceBriefings
Talking about e-science

Security and e-Science

'Password'; '123456'; '12345678'. The top three most popular passwords of 2012, as published in lists by hackers, were identical to the top three of 2011. When it comes to security, popular passwords aren't to be celebrated – not only are these passwords easy to guess, but it's a safe bet large a majority of easily guessed passwords in 2012 were 'protecting' the very same files they did in 2011. And, more often than not, those same passwords are also duplicated across a range of online services. This creates an easy target for identity thieves, whose intent is much worse than those who publish passwords online. What the fact that such lists can be published highlights is that total security is elusive: whether passwords are easy to guess or not, they are sometimes liberated, even from the biggest sites in social media and online gaming. Is the password reaching a crisis point? And what could replace them?



Popular passwords should be avoided at all cost

e-Science faces the same challenges as the rest of the online world – not least because many researchers are online outside of work, just like everybody else. But there are specific concerns: e-Health will herald a new era of personalised medicine, but having your personal file compromised could reveal more about you than you would like. Any large corpus of knowledge could be vulnerable to attack by cyberterrorists, and the hyperconnected 'smart cities' of the future might be an attractive target for acts of cyberwarfare. That is, if they're not quite smart enough to outsmart the bad guys.

e-infrastructures such as the grid also have a long history of managing security, access to services, and controlling privileges. These concepts are becoming more and more important to the rest of the online world, as the idea of universal 'web identities' takes hold. Indeed, multifactor authentication solutions ('two stage sign-ins') employed by sites such as Facebook and Google generate one time passwords to be provided in addition to your normal password.

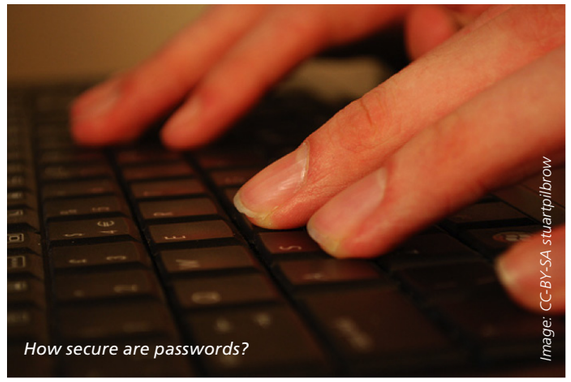
At the same time, e-science services are beginning to adopt social media credentials to allow users access. While this may simplify access from a new user's perspective, opening up grids and academic clouds to more researchers in the life sciences and e-Humanities, it can present new security challenges.

Stephan Lüders, CERN Security – *"Computer security is a sociological problem. It is time to teach our users and colleagues to stop-think-click when browsing the Internet as they've been taught to look both ways when crossing a road."*

Grid Security: Certificates

With many people becoming overwhelmed by the growing number of web-based services they use daily, the concept of a universal web identity seems like a sensible solution. For researchers using the grid, this idea is familiar in the form of digital certificates. Certificates are files that reside on a user's personal device that contain, alongside information such as date and duration of validity, a special key that is unique to the user, generated by a certificate authority. When a user accesses the grid, their credentials are checked to see if they are authorized to do so by the certification authority.

Certificates may be familiar to people outside of the grid community using services such as OpenVPN to securely access their employer's network. This typically requires a certificate issued by your employer to be installed; some proprietary Virtual Private Networks, such as that provided by Cisco, are configured to use your work login and password for authentication. Websites also use security certificates, but this may only become apparent to the average user when the certificate expires.



How secure are passwords? *Image: CC-BY-SA stuartpilbrow*

Certificates have some advantages. Because they identify an individual, losing or having a laptop stolen that is validated by having a certificate installed only requires a single certificate cancellation request be made. This stands in contrast to the many different logins and passwords the typical user has for web services, which would all have to be changed individually if a laptop was stolen – some browsers contain easily accessible lists of logins and passwords used, for example.

The idea of a single web identity, therefore, with a single sign-on seems to have some value in the rest of the online world. Several protocols including OpenID, OAuth Connect and Facebook Connect (a proprietary protocol) have arisen. The latter two, respectively, allow Twitter and Facebook credentials to be used to sign in to all kinds of web-based services. Due to the prevalence of social media these are becoming de facto standards, even for some grid services. There are concerns about the security risks of using social media as a universal sign on for e-science services (which

Roberto Barbera, Chain project – *"With the same simple sign on, a user could access everything from their campus network via Eduroam to the entire global grid. This is tremendously powerful"*

is plausible, because social media is such a huge target for online fraud). However, these "are removed by retaining a distinction between authentication and authorization," says Roberto Barbera, Technical Coordinator of the CHAIN project, who are providing access to grid services using social media authentication.

That authorization could come in the form of certificates or portable IDs such as the Shibboleth system from the UK organization, JISC, which is what CHAIN is using. The user can then access grid infrastructures around the world, including EGI, Open Science Grid, Teragrid, GISELA, SAGrid and Garuda. "You have to remember that authentication is completely separate from authorization," says Barbera, "identity federations allow us to control access, but we can control the privileges a user has separately." The report 'Advancing Technologies and Federating Communities' produced by Terena suggests that more researchers are using the social web to collaborate, and that e-science services should provide access via social credentials. However, a research institute needs to be sure that the individual presenting social credentials is the same as the individual they have authorized to use the service.

OpenID Connect

At the time of writing, OpenID Connect, a suite of lightweight universal ID standards, is in the implementer's draft phase. It aims to offer an alternative to Facebook Connect and OAuth and is said to be an improvement of the previous version of OpenID.

Clouds: The Safest Place?

It's a recurring story: an individual working for an organisation providing some sort of public service loses a USB key, or has their laptop stolen. The files contained on the stolen item contain the personal details of thousands, tens of thousands – or more – individuals. Worse, the data was unencrypted, meaning anyone could access it easily. Amidst public anger that their personal information could be so easily accessed, new security measures are put in place. Usually the solution chosen is encryption, so all data, whether sensitive or not, must be encrypted, whether on a laptop's hard drive or on a USB stick.

But does that solve the problem? It's a technological solution that's fairly easy to implement, but it ignores the fact that, in many cases, poor practices undermine its usefulness. Password-based encryption is only as effective as the strength of the password. And if the device is left logged in when not in use, it may as well not have been encrypted.

Clouds have the potential to offer much better security. Access can be controlled to files, to greater or lesser degrees for the individual collaborators working on those files, as required. Thanks to desktop synchronisation and version control (which saves incremental changes to files at many points in time over the lifetime of the file) storing everything in the cloud seems like the perfect solution. So much so that some netbooks only allow saving to the cloud. But which cloud?

Keeping data on USB keys is a risky business. *CC-BY-SA Scienceatlife (Flickr*

Popular cloud services such as Dropbox, Box.com and Google Drive have been widely adopted by researchers (and many others) because they offer easy ways to collaborate on or share files and data. The reliability and of these services is very high, but as with any online service there always remain vulnerabilities and potential for attack at every level (from a user's personal machine and network, to outsourced online helpdesks with lax security)[1].

As e-Health, which will offer unparalleled diagnosis capabilities and personalised therapies, becomes a reality, there is a need for clouds that operate independently of the privately run clouds. Stratuslab in Europe produces software that allows researchers to build academic clouds on their own hardware. These are not subject to the same terms (or potential for change of terms) as commercially available services. For researchers in many fields, a long-term goal is the establishing of repositories for 'Big Data' coming out of computationally intensive science that stand apart from the 'free' commercially run cloud services, with the different terms of service they entail.

Just as for jobs on the grid, controlling access is a key issue. Strong passwords, certificates, or logins tied to a machine's hardware all have their place.

1- Dropbox, for instance, had a major security flaw in 2011, and in February 2013, Zendesk, an online help system for Twitter, Tumblr and Pinterest, which also looks after Box.com support, experienced a security breach: http://www.wired.com/threatlevel/2013/02/twitter-tumblr-pinterest/

The Password Problem

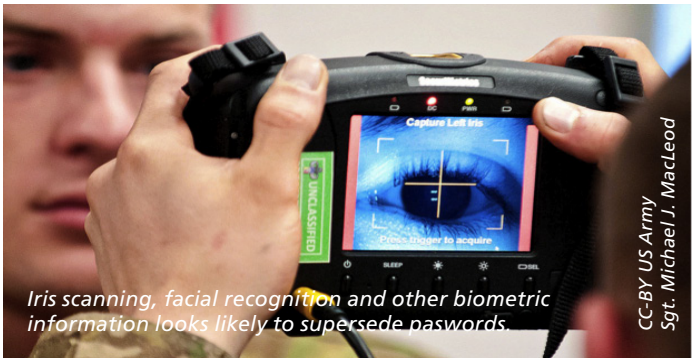
In spite of major efforts to educate users around the importance of using unique passwords, it seems that many users disregard warnings that their data is at risk of being compromised. According to reports given in several high-profile hacking cases involving attacks on state-level systems there are some serious short falls in security. Passwords are routinely distributed indiscriminately, rarely or never updated, even displayed on post-it notes in areas that can be physically accessed by individuals who would not otherwise have been given the password.
When it comes to judging the security of a password-protected system, a concept called Kerckhoff's principle is often invoked: a system should be secure even if everything about it, apart from the password, is public knowledge. For open standards advocates, this is a cornerstone of good encryption practice – the 'security by obscurity' used by proprietary software engineers avoids Kerchoff's principle precisely by not making the workings of the system public knowledge. However, if there are loopholes (and there often are), hackers can find them and circumvent any security measures in place. In open source software, the 'many eyes' working on the same code are more likely to spot loopholes, and many minds working on fixing the loopholes will do so more effectively.

But in practice, passwords in software, whether proprietary or open source, suffer the problem that people forget them. Many online services offer a 'I forgot my password' option at login, which will send a password reset code to your email address. While this might seem sensible it could be the first step to having your entire online life hijacked. Even services that offer extra layers of protection, such as requiring you to provide personal details, can be easily duped into sending a password reset to an unauthorised third party. Service providers have to manage high levels of password reset requests; people just have bad memories, so its unsurprising that they are so ready to be helpful in helping us get to our accounts.

Passwords are either so difficult to remember, therefore, that they have to be reset, or so easy to guess at that they provide no protection whatsoever. Perhaps passwords alone are no longer the answer. Two-step verification, now being adopted by services such as Google and Facebook, combines passwords with device-dependent ID keys that have to be set up for first time use. This is only secure if users set their devices to lock out during periods of non-use.

Sven Gabriel, NIKHEF – *"In a distributed environment like the European Grid Infrastructure, operational security has an additional dimension, since here we have to coordinate the activities of many different security teams involved in a multisite incident" See how distributed teams combatted a simulated virus on the grid at: http://v.gd/gridsecurity*

In 2012 researchers used the computing prowess of supercomputers to show that even the strongest password keys could be broken by brute force. The falling cost of computing power means that, in the long term, the age of the password is drawing to a close.
Biometric ID authentication based on facial or iris recognition looks likely to play some role in how we use devices in the future. Such technologies have been in place for a number of years at national border controls, and are becoming more commonplace in mobile devices. How web freedoms can be maintained in a future where our bodies become our logins and passwords is likely to be an area of intense discussion.



Iris scanning, facial recognition and other biometric information looks likely to supersede passwords. CC-BY US Army Sgt. Michael J. MacLeod

How secure is your platform?

The choice of hardware and software platforms used in working environments greatly affects their susceptibility to infection by malware, the safeguards against data loss, and their overall security. UNIX-derived operating systems favoured by the e-science community have historically tended to be more secure than proprietary alternatives like Microsoft Windows for the simple reason that the former offer greater granular control of access and editing privileges to files. UNIX-like systems including the many derivations ('distributions') of GNU/Linux and also Mac OS X (from FreeBSD/Mach) observe a clear distinction between users and administrator. Many of the security loopholes in Windows were historically caused by systems being installed in administrator mode by default, which lets malicious code be surreptitiously written to locations in the system without the owner's knowledge; many of these instances of code being written would have alerted users on UNIX-like systems by demanding a password. However, Microsoft's increasingly swift update cycle has improved security greatly in recent years.

Indeed, Microsoft's Windows is conspicuously absent from a recent list of top ten vulnerabilities produced by security software firm Kapersky, yet is still the most targeted platform, mainly due to market share. Linux is the most popular platform in e-science for other reasons that also benefit its security: as it is an open platform, 'many eyes' are involved in checking the code for loopholes. And as this is performed openly, it is the embodiment of Kerckhoff's principle. Additionally, the many different distributions allow users to tailor the platform for their research. This creates a diversity that is as sound a defence against virulent malware in the online world as it is in agriculture. Monocultures in plants or computers means threats can propagate quickly with devastating results.
Any operating system that grants application plugins such as Java or Flash special privileges can, however, compromise security. They effectively create a virtual monoculture. Examples of malware on various systems has subsequently led to the latest versions of the Google's Chrome and Mozilla's Firefox browsers being released with Java turned off by default. HTML5, an open standard, is being promoted as an alternative for developers building web-based services.

A growing trend for employees to 'BYOD' – bring your own device – also presents challenges that differ from those experienced before. BYOD presents challenges to organisations aware of the productivity boosts adopting such a policy can make, because it can introduce security risks depending on the device. Apple's iOS ecosystem is generally more secure than Android due to the fact that each submission is checked for content and functionality, leading to a closed app ecosystem. Users wishing to 'jailbreak' out of the so-called walled garden who have not subsequently protected their devices from unauthorised access can present security risks accidentally, but the Android ecosystem, where applications can be freely distributed without undergoing checks, presents the bigger risk.

Summary
e-Science faces the same challenges of authentication, universal identity management, and authorisation (including privileges) as many other web services. But with the number of researchers using such services in light of the growing importance placed on Big Data for life sciences and e-Health, for example, it is important that access to them is properly and securely controlled. The changing nature of how people use the web for the rest of their online lives is also influencing how people access e-science services, and it looks likely that universal web identities might prevail over the anonymity of the early days of the web.

For more information:

Stefan Lüders' blog post on passwords: ***http://security-blog.switch.ch/ 2013/01/09/password-awareness/***

Advancing Technologies and Federating Communities : ***http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/aaa-study-final-report.pdf***

CHAIN: ***https://www.chain-project.eu/***
OAuth: ***http://oauth.net/***

EGI : ***www.egi.eu***

Real Time Monitor: ***rtm.hep.ph.ic.ac.uk***

iSGTW: ***www.isgtw.org***

e-ScienceTalk: ***www.e-sciencetalk.org***

email: ***info@e-sciencetalk.org***

Scan this QR code into your smart phone for more on this e-ScienceBriefing

e-ScienceTalk is co-funded by the EC under FP7 INFSO-RI-260733