

Security & Privacy Challenges in E-Learning 2.0

Edgar Weippl, Martin Ebner
Secure Business Austria / Graz University of Technology
Austria
martin.ebner@tugraz.at; weippl@securityresearch.at

Abstract: E-Learning 2.0 uses Web 2.0 tools for e-learning. New services on the Internet can be swiftly integrated into existing applications; students can create MashUps, for instance, using a variety of services on the Internet. The main risk comes from the fact that students and teachers are not entirely aware that their institution does not control these services. The servers are located in a variety of countries, thus privacy laws also differ. In addition, as most Web applications are built as three-tier architectures, typical security weaknesses exist, such as invalid input, a lack of server-side checks, and excessive privileges.

Introduction

E-learning dates back to the correspondence-based learning of past centuries and has subsequently evolved to computer-based training, web-based training, and more recently to "E-Learning 2.0." The term e-learning 2.0 was coined for the first time by Stephen Downes (2005). He defines e-learning 2.0 as the "*use of Web 2.0 technologies in educational context*". In the present text we will discuss two aspects of this: Web 2.0 and the educational context.

1.1 Web 2.0

As predicted by a number of computer experts, the World Wide Web evolved into a large worldwide network in the early 1990s. Tim Berners-Lee (1989) mentioned that there is a necessity for a network that anyone from anywhere in the world can contribute to. However, for more than ten years active participation on the web was reserved to a small group of people who had the sufficient technical knowledge to use HTML or similar web-based languages. Web 2.0 is characterized by the interactive nature of applications, such as Wikis and Blogs, that allow users with little technical background to actively contribute content.

Today, the Web is used in a completely different way than it was ten years ago. Access to the Internet is widely available and end users can choose from a broad range of devices (PCs, laptops, PDA, cell phones, etc.). Based on these two conditions, improvements in usability of Web applications and research in Human-Computer Interaction help make the Web usable for a broader range of people. Wikis, Weblogs, and Podcasts are not really technological revolutions, but social ones. The necessary basic technology has been available for years, but has now been assembled to web-based software that is easily accessible and usable. There are, obviously, a great number of interesting research questions to address—for us, a field of particular interest is e-learning.

1.2 Educational Context

The second aspect mentioned by Downes is much more important within a discussion of e-learning 2.0. How can we use Web 2.0 technologies in educational settings? The use of high-end tools will not necessarily lead to improvements in the outcome of learning, just as technology for e-learning, in general, does not improve learning in terms of teaching (Moyer, 2002, Feuerstein, 1980)—improved teaching is made possible, in part, by e-learning technology, and it is this improved teaching that leads to improvements in the outcome of learning.

Learning is a basic social and cognitive process and must be performed by learners themselves. Neither computers nor Web-based applications are able to improve the outcome of learning, per se, but they can support the learning process by changing a lecture's didactical aspects, offering a variety of learning styles, and increasing

learners' motivation. In other words, to assist participants, the use of Web 2.0 technologies must be embedded within a concrete teaching and learning context.

The phenomenon Web 2.0 is best characterized by the sayings: "The user is the content", "Internet is about *people*", or "Give and take culture". Should a person want to take part, they can contribute and connect with the rest of the world. This is the first time that such an option is available in human history for reaching such a large number of people within a very short time frame. The interactive nature of Web 2.0 facilitates the unprecedented ease of authentic and meaningful learning activities (Oliver and Lake, 1997a, Oliver and Lake, 1997b, Herrington et al., 2003, Herrington et al., 2004) easier than ever before.

Mainstream research focuses on technical solutions, and—increasingly—pedagogical issues; privacy and security have not yet been adequately addressed. Considering the enormous costs involved in creating and maintaining courses, it is surprising that security has not yet been considered an important issue by all stakeholders; moreover, as the number of users increases and major portions of a student's learning history are recorded electronically, privacy requirements should become more relevant.

The major security topics for "traditional" e-learning that focused on the delivery of content were content protection and access control. Recently, the term "Web 2.0" was coined to reflect the changed nature of interaction. Web pages are no longer created by an individual and read by the public; instead, the "audience can create its own content" using interactive Web technologies, such as Blogs and Wikis.

In the past, academics were largely unconcerned about security, mainly because users in academic areas tended to be non-malicious. Today, however, campus-wide IT-security is crucial because e-learning applications and infrastructure have become "business"-critical applications. IT-security is usually addressed in a technical way by IT-departments; management does not (yet) see it as a part of the overall "business strategy" of a university. For example, almost all institutions install firewalls and anti-virus software to protect campus resources but fail to perform adequate security management.

Countermeasures are usually implemented to "fix" problems, such as blocking typical P2P ports if a university is at risk of being sued by the music industry. Only recently have universities recognized the importance of securing personal data in all departments. Social security numbers and credit card information are valuable assets used in identity theft. Attacks on these assets were successful, for instance, at the University of Colorado (Crecente, 2004). Similar incidents occurred at the University of Texas (Associated Press, 2004) and at the University of Colorado (Patel, 2007).

Even the most common security safeguards have drawbacks that are commonly overlooked. At Stanford University, for example, the residential computing office selected an anti-virus program. However, the program can be set to collect data that could potentially violate students' privacy expectations; therefore, many students decline its use (cf. Herbert, 2004)

1.3 Definitions Related to Security

While there are many definitions for the primary requirements of security, we will use the classical CIA requirements. CIA is the acronym for confidentiality, integrity, and availability. All other requirements can be traced back to these three basic properties. Avizienis et al., (2004) define confidentiality as the *absence of unauthorized disclosure of information*, integrity as *the absence of improper system alterations* and availability as *readiness for correct service*.

Dependability is a broader concept that encompasses all primary aspects of security save confidentiality, and in addition (1) Availability; (2) Reliability, which refers to the *continuity of correct service*; (3) Safety, which is defined as *the absence of catastrophic consequences on the user(s) and the environment*; (4) Integrity; and (5) Maintainability, which is *the ability to undergo modifications and repairs*.

In our context, security means that in a secure teaching environment, users need not be concerned with threats specific to e-learning platforms or to electronic communication in general. A secure learning platform should incorporate all aspects of security and make most technical details transparent to instructors and students. However, rendering a system "totally secure" is too ambitious a goal since no system can ever be totally secure and still remain usable at the same time.

2 E-Learning 2.0

Graz University of Technology has a lot of experience in the area of traditional e-learning (e.g. Maurer, 1996), and is also gathering research results for educational approaches concerning e-learning 2.0 (Ebner, 2007, Ebner et al., 2007, Schinagl, 2006). Some few examples are listed below:

E-Learning Blog: Since April 2006, all of the university's e-learning activities have been available on a Weblog. Reports, presentations, publications, and Podcasts give very detailed insight into the work of the Department Social Learning. This form for presenting the information was chosen not only for presentation purposes, but also for sharing content with peers working in similar fields. Moreover, the possibility to comment on each of the contributions facilitates discussion and "opens" the university to the rest of the world.

TU Graz LearnLand: In October 2006, a so-called Blogosphere was launched at the University. The software is based on the Open Source product ELGG. Lecturers and students can create their own Weblogs through a simple Logon. Further features are provided automatically, for example, embedding videos or pictures and a social book marking system (with import and export) and tag clouds. In March 2007, Weblogs were used in special didactical settings. By integrating them into daily lectures, researchers collect in-depth experience on how this technology works.

Bauwiki: One of the first attempts to use a Wiki in Education took place in 2005. The free TWiki System (<http://www.twiki.org>) was installed to support lectures in the field of Civil Engineering. A number of studies were carried out to show how such a system works in real life scenarios and educational settings. In sum, Wiki systems support all kinds of group activities and online collaboration. Nonetheless, it is necessary to also discuss control mechanisms for students' contributions. Needless to say, when an article is finished, it is difficult to distinguish how much has been done by each participant. The risk of allowing freeloaders to pass through the system is immanent and constant. This might lead to a completely different view of Wikis and their potential: Wikis should help increase cooperation between various people. Students should learn how other opinions fit with their own, reflect on various standpoints, and in the end come to some agreement with each other. Wikis therefore seem appropriate, but they also enforce a new didactical concept. The effort of each learner is no longer significant; instead, lecturers have to evaluate the results of the group. Teamwork is often postulated as necessary in education and Wiki can support this by granting the individual effort secondary importance.

3 Security in E-Learning 2.0

For many universities, e-learning systems have become assets critical to production. It is thus imperative to evaluate all of the generic requirements (confidentiality, integrity, and availability) during a process of risk assessment (Weippl, 2005c). The first step in such a process is to understand security as including all factors enabling technology. Only when systems work reliably will users trust them and use them (Weippl, 2001).

Learning is a creative process and learners should use different learning materials in their learning processes. Web 2.0 technologies help to link this material together, share, and expand it in a simple way. The major difference with E-Learning 2.0 is that complex and interdependent Web applications are used. For instance, MashUps can be created by combining videos (e.g., from You Tube), pictures (for example from Flickr), presentations (e.g., from SlideShare), and podcasts (e.g., from iTunes). By searching the Web, incidental learning occurs even through reading or hearing a great number of different statements.

However, there are three potential risks. First, complex applications raise the likelihood of vulnerabilities due to design and coding errors. Second, it is difficult to know how much work students plagiarized, that is, copied and linked from other sources without giving due credit. Third, publicly available sites may offer too much information about students and their work. Without the privacy of the classroom, students may either refrain from certain things (e.g., nude photography in art classes) or may later regret having published them.

3.1 Weak Web Applications

Typical weaknesses of Web-based applications are caused by (1) invalid input, (2) missing server-side checks, and (3) excessive privileges on the server.

Distributed applications are vulnerable to attacks by *unvalidated input* data. Sending requests that are too large, that have illegal characters, or simply contain other unexpected input may cause server-side applications to

behave in unspecified ways or crash. Countermeasures include the thorough implementation of the Clark-Wilson security model (Clark and Wilson, 1987), systematic testing and fuzzing.

Web application programmers often assume that users use only browsers to interact with their server-based Web applications and would never send http requests directly using, for instance, wget or modify requests with a local proxy. While *client-side checks* are certainly useful to improve the user interface, they must not replace server-side checks. Requests to the Web server can be modified arbitrarily and all input must be checked by the server-based application prior to processing.

Too often, processes are executed with *excessive privileges*. Web applications usually connect to the database with a single user who also has too many privileges. A simple error in the application can compromise the security of all data stored in the database. Typical attacks are SQL injection and cross-site scripting. A defense-in-depth approach can help to contain vulnerabilities. If, for instance, the application can only execute stored procedures in the database to perform the required operations and cannot directly access the tables in the database, the probability of successful SQL injection attacks is clearly reduced.

Most e-learning applications are built as three-tier applications, comprising a data tier, an application tier, and a client tier. Moodle, for instance, uses a MySQL or PostgreSQL database on the data tier. The e-learning application (application tier) is Moodle itself. It is written in PHP and executed by the Web server—in most cases, an Apache Web server. The client tier is the Web browser used by students and teachers to access Moodle.

Currently, e-learning application security focuses mainly on the application layer. Moodle developers, for instance, maintain a dedicated site on security issues pertaining to Moodle. On the application level, Moodle—like most other platforms—seems to adhere to practices of secure coding (Howard and LeBlanc, 2002). Therefore, typical vulnerabilities at the application level, such as buffer overflows and unchecked input, are less likely to occur.

However, even if the application itself is secure, a variety of other attack paths remain. Unauthorized users might gain access to the Web server, the PHP module, or the operating system of this server. They can then modify the code of the application and subsequently alter any data in the data store because the e-learning application has full access to data storage. Moreover, the log files in Moodle are created at an application level only. Therefore, log messages can be manipulated by the altered application.

In addition, an attacker can also directly attack data storage, i.e., the database management system or the operating system of the database server. Even if safeguards, such as triggers, were installed in the database, the underlying files and log entries could still be modified. Many Moodle installations use MySQL as their database. The problem with MySQL is that it does not support many of the more sophisticated access control mechanisms offered by PostgreSQL, Microsoft's SQL Server 2000, or Oracle 10g. Clearly, attackers can and will look for the easiest way to attack a system. Given the many potential vulnerabilities of a default three-tier installation, one must be cautious when trusting data stored in the system.

3.2 Plagiarism

Currently, a variety of anti-plagiarism software exists on the market with the aim of determining whether authors used other sources and included verbatim copies of intellectual property. This topic gains even greater relevance with the hype of Web 2.0, in particular, and the possibility of linking, sharing, and embedding sources from other websites (Maurer et al., 2006). The “copy and past” phenomenon leads to a rethinking of control mechanisms in educational settings. Lecturers must be aware of methods to assure that work handed in really comes from the student.

While anti-plagiarism software may certainly help to find sources on the Web in digital libraries, it cannot determine whether a student actually plagiarized. Manual checking is required to determine if texts highlighted for plagiarism are cited correctly. *Deep linking*¹ other people's Web content in one's own work may violate their usage policies or copyright and can constitute plagiarism if not properly acknowledged.

3.3 Publicity versus Privacy

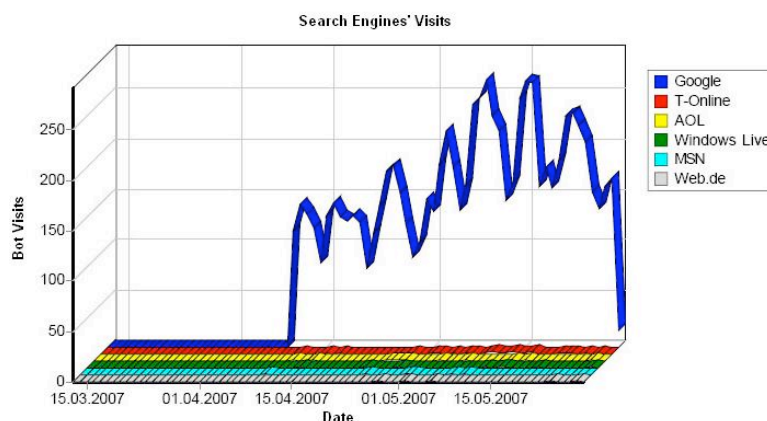
Nowadays, it is necessary to point out that the majority of typical Internet users are not aware of the processes working in the background of Web 2.0 applications. Currently, the effects of leaving digital traces may be considered unimportant, but there is no guarantee that this will be the case in the future. Nowadays, we write our

¹This occurs when a user is taken from a hyperlink deep into the structure of a web site. This is considered to be bad practice as website structure may change with time and the link may become broken or inappropriate. (quoted from http://sensacom.com/web_glossary.html, last visited Feb 20, 2008.)

statements in Weblogs, contribute to a Wiki, and read/send our e-Mails via public providers. Currently, nobody knows or cares that other people can find this data when searching. The danger to today's "digital students" is that they are digitally recorded and copies are irrevocably stored in archives, such as archive.org. The mandatory use of services on the Internet for teaching, in particular, must be seen as critical since personal data is stored on servers that do not belong to the educational institutions.

Shown below are the statistical data from TU Graz LearnLand concerning the crawling of search engines. Without any additional measures, Google Bots, for example, hit the Web 2.0 application 200 times per day on average. By understanding the crawling process it becomes clear why so called "social software" applications get such a good ranking in Google.

Each contribution of a Weblog is indexed and since the entire site changes with every new blog entry, crawlers tend to visit active blogs quite often. In other words, if a Weblog is steadily updated and constant new contributions appear, bots will find it and increase the rank.



Neal, (2004) emphasizes that using new systems leads to new security and privacy issues. For example, wireless transponders used for toll collection on roads also create electronic traces of drivers. Similarly, online time, participation rates, and reaction times of students are recorded by e-learning systems.

Chadwick et al. (2003) present opposing views on privacy and defines related terms, such as pseudonyms, aliases, nicknames, anonymity, privacy, and confidentiality

"Anonymity ensures that others cannot determine your true identity. An anonymous person effectively does not have an identity. A pseudonym on the other hand is an alternative (fictitious or assumed) identity for a person. Aliases are identical to pseudonyms, but the two are used with different connotations in different contexts. ... Pseudonyms and aliases will usually not prevent the true identity of a person from being determined, although it may be difficult and may require law enforcement to enable it. Anonymity on the other hand should ensure that the true identity of a person is never found out, nor is capable of being found out. A nickname is also an alternative name for a person, and is often chosen as a friendlier variant of the person's formal name. Nicknames are not chosen so that the person's identity can be hidden."

The importance of privacy considerations is reflected by recent laws, such as the Patriot Act in the US and the Rip Act in the UK. The Patriot Act allows "a much broader category of cases to use information developed under the Foreign Intelligence Surveillance Act, where those subject to wiretaps are not informed of the surveillance even after the fact (Swire and Steinfeld, 2002)." In the UK, a comparable law, the Rip Act, has been passed. Some provisions require that systems must allow law enforcement agencies access to data, and that keys be provided to decrypt encrypted content. However, there seems to be no legal requirement to keep all data in long term archives. Thus, the privacy of individuals can be increased in the long term by regularly reusing backup tapes so that old content is replaced.

4 Conclusion

The use of Web 2.0 applications is increasing rapidly and new applications are available almost daily. Needless to say, from an educational point of view, the possibilities of computer-supported education may seem unlimited. Novel technology empowers teachers and learners to be creative in new ways. New didactical concepts

can be experimented with: for instance, MashUps enrich lecturing and learning by providing learning material from different sources.

Nonetheless, there remains the issue of privacy remains a problem to be contended with. Students store their content on different servers that are not under the school's control. The servers are, moreover, most likely located in various countries, which means that laws will differ and there is no way of efficiently addressing privacy issues in different jurisdictions. The dynamic nature of Web 2.0, however, requires schools to permit the use of multiple services and applications.

References

- Associated Press (2004). Former student indicted in computer hacking. USA Today. http://www.usatoday.com/tech/news/computersecurity/hacking/2004-11-05-u%t-hack-charge_x.htm.
- Avizienis, A., Laprie, J.-C., Randell, B., and Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions of Dependable and Secure Computing*, 1(1):11–33.
- Berners-Lee, T. (1989). Information management: A proposal. *Berners-Lee1989_InformationManagement*.
- Chadwick, D., Olivier, M., Samarati, P., Sharpston, E., and Thuraishingham, B. (2003). Privacy and civil liberties. In Gudes, E. and Sheno, S., editors, *Research Directions in Database and Application Security*, pages 331–346, Kings College, Cambridge, UK. Kluwer.
- Clark, D. and Wilson, D. (1987). A comparison of commercial and military computer security policies. In *IEEE Symposium on Security and Privacy*, page 184.
- Crecente, B. D. (2004). Hacker breaks into computer at cu. *RockyMountainNews.com*. http://rockymountainnews.com/drmn/local/article/0%2C1299%2CDRMN_15_3300%285%2C00.html.
- Downes, S. (2005). E-learning 2.0. *eLearn*, 2005(10):1. <http://doi.acm.org/10.1145/1104966.1104968>.
- Ebner, M. (2007). E-learning 2.0 = e-learning 1.0 + web 2.0? In *Proceedings of the 2nd Conference on Availability, Reliability and Security (ARES'07)*, pages 1235–1239, Vienna. IEEE Computer Society Press.
- Ebner, M., Holzinger, A., and Maurer, H. (2007). Web 2.0 technology: Future interfaces for technology enhanced learning? In Stephanidis, C., editor, *Proceedings of the 12th International Conference on Human-Computer Interaction (HCI 2007)*, volume 7, Beijing. Springer.
- Feuerstein, R. (1980). *Instrumental Enrichment: An Intervention Program for Cognitive Modifiability*. University Park Press, Baltimore.
- Herbert, D. (2004). Bigfix may be big risk, say rccs. *The Stanford Daily*. http://daily.stanford.edu/tempo?page=content&id=15170&repository=0001%_article.
- Herrington, J., Oliver, R., and Reeves, T. (2003). Patterns of engagement in authentic learning environments. *Australian Journal of Educational Technology*, 19(1):59–71.
- Herrington, J., Reeves, T., Oliver, R., and Woo, Y. (2004). Designing authentic activities in web-based courses. *Journal of Computing and Higher Education*, 16(1):3–29.
- Howard, M. and LeBlanc, D. (2002). *Writing Secure Code*. Microsoft Press, 2nd edition edition.
- Maurer, H. (1996). *HyperWave - The Next Generation Web Solution*. Addison-Wesley Longman Publishing Company, London.
- Maurer, H., Kappe, F., and Zaka, B. (2006). Plagiarism—a survey. *Journal of Universal Computer Science (JUCS)*, 12:1050–1084.
- Moyer, L. G. (2002). Is digital learning effective in the workplace? *eLearn*, 2002(5):5.
- Neal, L. (2004). Expectations of privacy. *ACM eLearn Magazine*. http://www.elearnmag.org/subpage/sub_page.cfm?article_pk=9344&page_num%ber_nb=1&title=COLUMN.
- Oliver, R. and Lake, M. (1997a). Training teachers for distance education programs: Using authentic and meaningful contexts. *International Journal of Educational Telecommunications*, 4(2):147–179.
- Oliver, R. and Lake, M. (1997b). Training teachers for rural and distant education: Using authentic and meaningful contexts. In Muldner, T. and Reeves, T., editors, *Proceedings of ED-MEDIA 1997*, pages 818–823, Charlottesville, VA. AACE.
- Patel, V. (2007). Hacker exposes info on cu students. *denverpost.com*. http://www.denverpost.com/sports/ci_5962767.
- Schinagl, H. M. W. (2006). Wikis and other e-communities are changing the web. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (ED-MEDIA) 2006*, pages 2858–2866. AACE.
- Swire, P. and Steinfeld, L. (2002). Security and privacy after september 11: the health care example. In *Proceedings of the 12th annual conference on Computers, freedom and privacy*, pages 1–13, San Francisco, California. ACM Press.
- Weippl, E. (2001). The transition from e-commerce to m-commerce: Why security should be the enabling technology. *Journal of Information Technology Theory and Application (JITTA)*, 3(4):17–19. http://peffers.net/journal/volume3_4/ecpreface.pdf.
- Weippl, E. (2005a). Non-repudiation and audits in e-learning, invited paper. In *Proceedings of E-Learn 2005*, pages 1785–1790, Vancouver, Canada. AACE.
- Weippl, E. R. (2005b). Dependability in e-assessment. In *Proceedings of ED-MEDIA 2005*, Montreal, Canada. AACE.
- Weippl, E. R. (2005c). Security in e-learning. *ACM ELearn Magazine*. http://www.elearnmag.org/subpage/sub_page.cfm?section=4&list_item=19&pa%ge=1.