LIBER QUARTERLY
The Journal of the Association of European Research Libraries
Vol. 27, no. 1 (2017) 45–57 | e-ISSN: 2213-056X

Web Interface Security Vulnerabilities of Selected European Open-access Academic Repositories

Matus Formanek

Department of Mediamatics and Cultural Heritage, University of Zilina, Slovak Republic
matus.formanek@mediamatika.sk

Martin Zaborsky

Department of Mediamatics and Cultural Heritage, University of Zilina, Slovak Republic
martin.zaborsky@mediamatika.sk

Abstract

The given analysis summarizes the status quo of the level of security of web interfaces of selected European academic repositories in the field of library and information science. It focuses on the presence and qualities of the secure HTTPS protocol via SSL/TLS protocols. The security of the transmitted data is particularly important in the network environment of the Internet, especially if log-in user data is transmitted. Disclosure may have a direct impact on saved digital objects and their metadata which together represent the most valuable parts of systems of digital libraries and repositories. Furthermore, the paper points to the most noticeable vulnerabilities of protocols of web interfaces and presents practical recommendations for the expert public. These may contribute to the increase of the level of security of the discussed systems. The authors base their proposals on the currently available scientific publications and scientific articles about the given topic.

Key Words: digital library; web vulnerabilities; security; repository administration

Uopen Journals | http://liberquarterly.eu/ | DOI: 10.18352/lq.10178

1. Reasons for Research Realization

Nowadays, the issues of computer security are of great interest, especially in a network environment where various electronic systems are closely interconnected and transmit sensitive user data. It is said that we live in a so-called Information Age, when "information security and privacy are very important issues" (Al-Suqri & Akomolafe-Fatuyi, 2012).

According to the current Internet Security Threat Report published by Symantec (2016, p. 18), *"effective security requires layers of security built into devices and the infrastructure that manages them, including authentication, code signing, and on-device security… Analytics, auditing, and alerting are also key to understanding the nature of threats emerging in this area. Finally, strong SSL/TLS encryption technology plays a crucial role in authentication and data protection."* There is much more information about the cybersecurity challenges and the threats in this online accessible document.[1]

ICT tools offer wide possibilities of access to information, which users may use quickly and simply thanks to the Internet. The aspect of network connection, however, brings certain risks. The most sensitive aspect is the system and network security of important electronic systems such as digital libraries or academic institutional repositories. They may fulfil their role only if they can offer their service securely and stably in the broader environment of the Internet, where *"information security and privacy are very important issues. Standards and mechanisms for the protection of the information during data transfer are also very important as technology changes and improvements in information storage are made, earlier information resources in print format need to be transferred to progressively newer technologies over time, as older forms gradually become obsolete"* (Al-Suqri & Akomolafe-Fatuyi, 2012). This results in such a state that *"the velocity of information dissemination sometimes overshadows integrity, the regulations and policies that govern the circulation of information"* (Al-Suqri & Akomolafe-Fatuyi, 2012).

Several experts have already studied these issues. In her paper, Kuzma (2010) dealt with the analysis of the security of web portals of 80 selected digital libraries in four European countries. She presents attacks of hackers on academic digital libraries in the American state of Indiana in 2002 and 2004 as a warning.

Increased security is one of the factors that may significantly increase the general value of network applications. It may also help to achieve a higher level of trust in online services (Chen, Choo, & Chow, 2006). The loss of trust of users may have harmful consequences in addition to the risk of personal information theft (Kuzma, 2010).

Web 2.0 tools offer possibilities for using user identity in the network environment due to which transfer of corresponding log-in and other data to various web applications (phone numbers, addresses, numbers of payment cards etc.) occurs. Web 2.0 applications run in browsers, which are mediators between users and applications. Various web threats have a higher impact nowadays than ever before (Šilić, Krolo, & Delač, 2010). All applications using networks—among which we count digital libraries as well—require a properly chosen security mechanism of the transferred data because these systems use and store the user credentials, satisfy information needs and should stay accessible online 24/7.

Studies about the issues of security of digital libraries, their web interfaces or content, are not easily available, especially in a required width and depth of topical take appropriate for the community of library workers. Kuzma (2010) also points out this fact. She claims that the issue of security of user interfaces of digital libraries is not studied enough, not even today. The expert community lacks literature about this topic. As a result, there is a lower level of awareness about security risks, which loom over all computer systems in the Internet environment, including repositories and digital libraries.

Another problem lies in the fact that library workers and librarians themselves do not often realize the aspects of computer security of the library systems and networks they work with (Fox, 2006). Fox (2006) adds that digital content is usually very valuable and library workers have to protect it as well as they protect data about visitors. We realize that these workers are not specialized in the field of security of IT systems but these problems must not be underestimated.

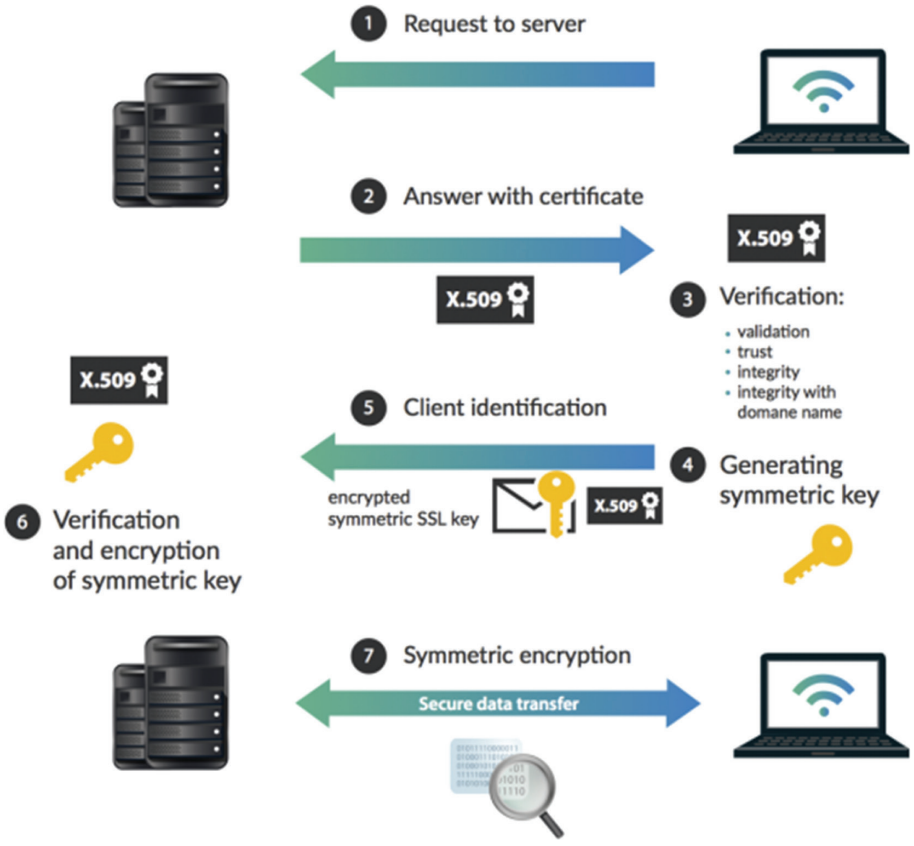
2. SSL/TLS protocols

One of the key requirements in secure communication within the network is the encoding of the connection to ensure that the communication is not

compromised, not even if the message is captured. For this purpose, security network protocols SSL (Secure Socket Layer) and the newer TLS (Transport Layer Security) are used. Technologies based on these protocols enable creating an encrypted connection between a client (e.g. a browser) and a server.

During authentication of a connection and transfer of messages, the SSL protocol uses the combination of symmetric and asymmetric encryption. As seen from the Figure 1 below, upon request for a secure connection, the server sends its public key together with a digital certificate (handshake stage) to the client. The client verifies the X.509 certificate's validity and if there are no doubts about the server identity, it generates a random number as a base for a

Fig. 1: The SSL technology principle (Internetum, 2015).



session key, which is encrypted by the public key of the server and sent back to the server. Using its private key, the server decrypts the obtained data and both sides create a unique session key. After the handshake stage, all communication is encoded using the created shared key, which is valid for the given session only (Rouse, 2014). When taking the ISO OSI reference model into consideration, the SSL protocol can be found in the presentation layer. In the TCP/IP model, it sits in the application layer (Kozierok, 2005).

The digital certificate of a server should be released and signed by a generally accepted certificate authority in order to prevent a scenario in which an attacker would pretend to be the certificate authority. Via visual clues, modern browsers inform the user that the communication with the web server is secure and it has a valid safety/security certificate issued by the accepted certification authority. The TLS protocol is a newer version of SSL and is compatible with it. For both protocols, the general name SSL or rather SSL/TLS is used.

3. Aims of the Research

The main aim of our research is to present the status quo regarding the use of secured web protocols by selected institutional repositories in Europe, to point out possible weak spots, and to propose recommendations for improvement. We analysed the LIS institutional repositories only because it fits the particular scope of our academic research about repositories. We want to adapt an academic repository based on open-source software for the LIS department of University of Zilina in the near future.

We do not want to point out security flaws of specific institutions (and expose them to attacks) nor decrease their status because administrators of digital repositories are not necessarily computer experts.

4. Methods

We chose repositories by selecting them from OpenDOAR.org—the authoritative directory of registered academic open-access repositories (University of Nottingham, 2014). We used the following selection criteria:

- they should be institutional repositories,
- part of the content of repositories as well as their interfaces must be available also in English,
- the repositories should be located in Europe,
- their focus must include the LIS area (Library and Information Science).

As of June 9, 2016, based on the aforementioned criteria, the openDOAR.org registry listed only 33 repositories. We studied all these repositories in more detail. We carried out the testing of the web interfaces in two stages:

Stage 1. Using the information stated in the record of every repository in openDOAR.org, we found a link to the main official page of the each system. Using the Mozilla Firefox v47 browser, we observed whether the interface natively supported the secured protocol HTTPS, either on the whole site or on the log-in page.

Stage 2. We tested those web interfaces that supported the HTTPS protocol using two independent tools:

a. SSL server test by Qualis SSL Labs company.[2] *"This free online service performs a deep analysis of the configuration of any SSL web server"* (Qualis SSL Labs, 2016).
The test focuses on the depth analysis of the current configuration of security certificates and supported cipher algorithms. It looks for vulnerabilities in the form of support of outdated technologies. Furthermore, the test simulates a so-called handshake of various versions of operating systems, browsers (Android, IE v6-11, EDGE, Firefox, Safari etc.), and JAVA web technologies.
b. SSL/TLS server test by High-Tech Bridge company.[3] It is "*aimed to enable anyone to assess how secure and reliable his or her SSL/TLS connection to a server (on any port) is, the service performs four distinct tests: Test for compliance with NIST Guidelines, for compliance with PCI DSS Requirements, for the most recent SSL/TLS vulnerabilities and test for insecure third-party content that may expose user's privacy"* (High-Tech Bridge, 2016).

We chose these tools because they represent a simple, available and mainly transparent way of testing the safety components of web pages. Tests may

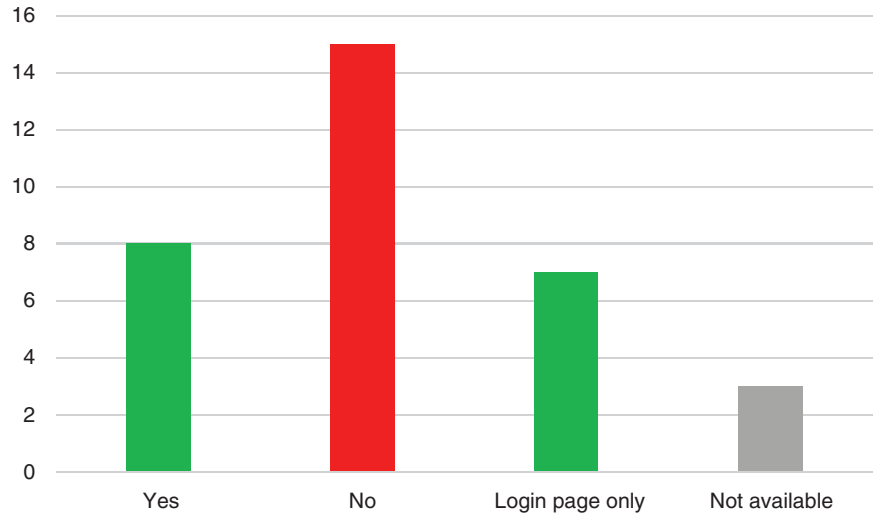
be carried out from any place and they may be repeated at any time. Another crucial element was the identical form of the results that these tests offer—they evaluate web portals using the usual scale from A to F, which is also used in the academic research. Partial steps, such as A- or B+ are also used to achieve finer granularity of the results. A+ represents a better level of evaluation than A which is better than A- and so on.

5. Test Results

During the first stage, 3 out of the 33 repositories did not work and thus it was impossible to determine whether their interface supported the secured web protocol. Their web domain was repeatedly unavailable during our analysis (May–June 2016).

As seen from the Figure 2 bellow, only 8 repositories use the HTTPS protocol natively in the whole interface (all web pages of a particular domain). That means that the encoded data transfer is available right after the user visits the page, or rather there is an automatic redirecting of the visitor from the unsecured HTTP protocol to the secured HTTPS (S = secure). The most important

Fig. 2: Use of the HTTPS in repository web interfaces.



transfer of sensitive user data (such as access names and passwords) occurs in the forms of login pages.

Many system administrators realize this fact and as a result only this login page is secured by the HTTPS protocol in exactly 7 repositories of our study.

The amount of studied academic repositories that do not use a secured protocol altogether, not even at login pages where users (as well as administrators) input their user credentials, is alarming: 15 systems (which is 50% of functioning systems involved in tests) **do not use any form of security for data transfer**. Those are not systems in the early stage of development or testing but production systems containing digital objects and content that is valuable for the given institution. This is unsettling because these systems are exposed to potential cyber-attacks and compromising of saved data.

In the second stage of our study, we looked into 15 web interfaces of individual repositories (8 in the web pages of the whole domain + 7 in the login page only), which used the HTTPS secured protocol. Again, test results were quite unsettling although we appreciate the use of the secured protocol.

In Figure 3, we see the scores that the SSL/TLS certificates obtained in two mutually independent security tests. The unsatisfactory level of evaluation F (Failed) is worth noticing because it was assigned to quite a large number of certificates of web interfaces of studied academic repositories. Three out of the total number of 15 certificates had the worst evaluation F in both tests.

We do not want to point out the flaws of particular web interfaces and draw negative attention to specific institutions. This is why we will not write the names of repositories nor institutions that cover them. The aim of the carried-out measurements is to point out the extent of the security problem, which needs to be solved in time. Having tried to help with the solution, we examined the most frequent causes for negative evaluations in the tests of SSL/TLS certificates in more detail. Figure 4 shows the most frequent reasons for bad scores. The substeps of the grades (such as A+, A- etc.) are not included in the Figure 4 because there are only small differences between the main grade and related substeps. The D grade is omitted because it was not obtained during the testing at all.

During the testing, we found that one of the most significant and currently highly-discussed security problems is a vulnerability known as

Fig. 3: Summarized scores for certificates.

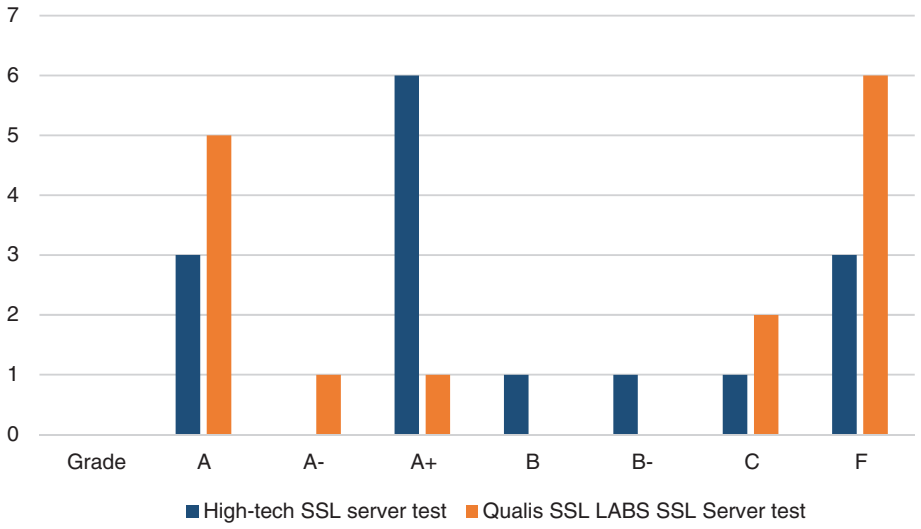
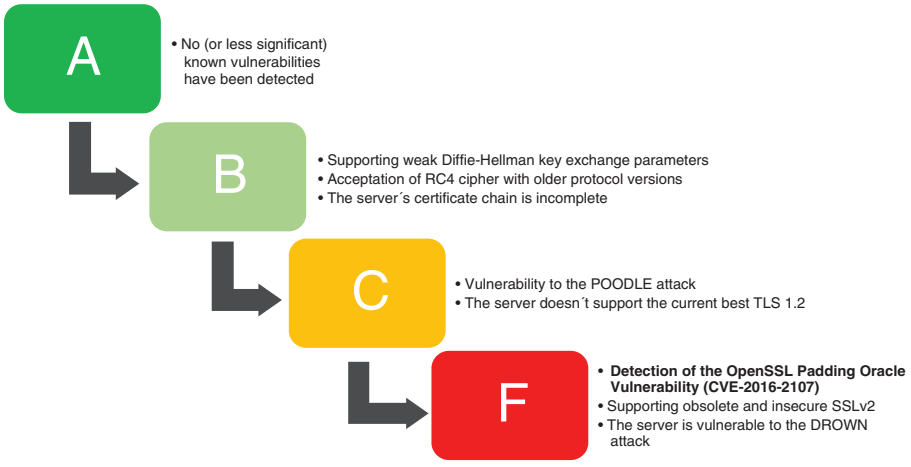


Fig. 4: Main reasons for low scores.



CVE-2016-2107 discovered by Juraj Somorovsky on 13 April 2016. It is the *OpenSSL Padding oracle vulnerability* as can be seen in Figure 4. The core of this problem lies in the fact that a so-called Man-in-the-middle *"attacker can use a*

padding oracle attack to decrypt traffic when the connection uses an AES CBC cipher and the server support AES-NI" (OpenSSL, 2016a).

Simply put, this vulnerability enables the attacker to decode the communication between the client and the server during the data transfer despite this channel being coded by means of the AES-NI algorithm (abbrev. from Advanced Encryption Standard New Instructions). Shortly after that, another vulnerability was discovered which is caused by a critical error in memory (using so-called buffer underrun). It is known as CVE-2016-2108. Both vulnerabilities were fixed in quite a short time. Afterwards, new fixed updates of cryptography and the SSL/TLS toolkit "OpenSSL" were released. Regarding both aforementioned vulnerabilities, the only secure protection against them is an update and use of the newest version of the toolkit. During the writing of this article, it was version 1.0.2h or 1.0.1t released on 3 May 2016 (OpenSSL, 2016b).

Two other aforementioned critical vulnerabilities, which caused the F grade, are caused by the server support of the out-dated SSLv2 protocol, which is not recommended for use (not even its updated SSLv3 version). According to expert security portals (such as disablessl3.com, digicert.com and others), it is customary to block the support of SSL v2/v3 protocols in browsers for the sake of security and because of the number of possible threats. Progressively, these protocols are replaced by much safer TLS protocols (the newest version is TLS 1.2). This step significantly prevents other vulnerabilities (which caused the C grades in Figure 4): it is "*a proactive way to combat the "POODLE" vulnerability*" (Digicert, 2016). When using TLS protocols, one must pay attention to their up-to-dateness and replace TLS 1.0/TLS 1.1 by TLS 1.2 wherever it is possible.

We did not describe rare or less serious vulnerabilities which are outside the scope of this article. More information about the topic may be found on specialized web pages, such as open-source project called OpenSSL (2016c), specialized webpages and technical papers about DROWN attack (Aviram et al., 2016) or about Diffie-Hellman key exchange (Adrian et al., 2015).

6. Conclusion

In this paper, we presented alarming results of our analysis. 50% of investigated European digital repositories (listed in OpenDOAR.org registry) that cover the field of library and information science **do not use any kind of**

transfer security for access and other user data. The relatively high number of bad scores (especially score "F" in any of the tests) for the certificates is alarming, too. We realize that we cannot expect expert knowledge about security of web interfaces from librarians. However, these people often administer digital libraries and repositories.

Rapid development in the field of network computer security and constant discoveries of new security vulnerabilities require repository admins to closely cooperate with IT experts in the field of web security. These will follow the newest trends, implement, and last but not least, update web certificates of online repositories. We think that in the university environment where academic repositories are found, it will not be a problem to secure appropriate technical support. Our recommendation for repository administrators is to use high-quality updated TLS 1.2 security protocols which protect the flow of sensitive users as well as admin access (and other) data. Many cyber attackers wait for just a little mistake unintentionally made by admins or users. When using security cryptographic web protocols, repository admins protect the online identity of the repository and its reputation as well as the personal data of users and visitors, and valuable digital objects found in the repositories.

References

Adrian, D., Bhargavan, K., Durumeric, Z., Gaudry, P., Green, M., Halderman, J.A., …, & Zimmermann, P. (2015). Imperfect forward secrecy: How Diffie-Hellman fails in practice. In *Proceedings of the 22nd ACM SIGSAC Conference on computer and communications security* (pp. 5–17). New York: ACM Digital Library. https://doi.org/10.1145/2810103.2813707. Retrieved October 7, 2016, from https://weakdh.org.

Al-Suqri, M.N., & Akomolafe-Fatuyi, E. (2012). Security and privacy in digital libraries: challenges, opportunities and prospects. *International Journal of Digital Library Systems, 3*(4), 54–61. https://doi.org/10.4018/ijdls.2012100103.

Aviram, N., Schinzel, S., Somorovsky, J., Heninger, N., Dankel, M., Steube, J., …, & Shavitt, Y. (2016). DROWN: Breaking TLS using SSLv2. In *Proceedings of the 25th USENIX Security Symposium* (pp. 689–706). Berkeley, CA: The USENIX Association. Retrieved October 7, 2016, from https://drownattack.com/drown-attack-paper.pdf.

Chen, S., Choo, C., & Chow, R.Y. (2006). Internet security: A novel role/Object-based access control for digital libraries. *Journal of Organizational Computing and Electronic Commerce, 16*(2), 87–103. https://doi.org/10.1207/s15327744joce1602_1.

DigiCert. (2016). *Disabling browser support for the SSL 3.0 protocol*. Retrieved June 5, 2016, from https://www.digicert.com/ssl-support/disabling-browser-support-ssl-v3.htm.

Fox, R. (2006). Digital libraries: The systems analysis perspective, vandals at the gates. *OCLC Systems & Services: International digital library perspectives, 22*(4), 249–255. https://doi.org/10.1108/10650750610706961.

High-Tech Bridge. (2016). *Free SSL server test: About the service*. Retrieved June 12, 2016, from https://www.htbridge.com/ssl/#about.

Internetum. (2015). *What is SSL certificate and how it works?* Retrieved November 23, 2016, from https://www.internetum.com/what-is-ssl-certificate-and-how-it-works/.

Kozierok, Ch.M. (2005). *The TCP/IP guide: A comprehensive, illustrated internet protocols reference*. San Francisco: No Starch Press.

Kuzma, J. (2010). European digital libraries: Web security vulnerabilities. *Library Hi Tech, 28*(3), 402–413. https://doi.org/10.1108/07378831011076657.

OpenSSL. (2016a). Padding oracle in AES-NI CBC MAC check (CVE-2016-2107). *OpenSSL Security Advisory (3rd May 2016)*. Retrieved June 5, 2016, from https://www.openssl.org/news/secadv/20160503.txt.

OpenSSL. (2016b). *Cryptography and SSL/TLS toolkit: Downloads*. Retrieved June 10, 2016, from https://www.openssl.org/source.

OpenSSL. (2016c). *Cryptography and SSL/TLS toolkit: Vulnerabilities.* Retrieved October 7, 2016, from https://www.openssl.org/news/vulnerabilities.html.

Qualis SSL Labs. (2016). *SSL server test*. Retrieved June 12, 2016, from https://www.ssllabs.com/ssltest.

Rouse, M. (2014). Secure sockets layer (SSL). *TechTarget Search Security*. Retrieved June 10, 2016, from http://searchsecurity.techtarget.com/definition/Secure-Sockets-Layer-SSL.

Symantec. (2016). *ISTR: Internet Security Threat Report, vol 21*. Retrieved November 23, 2016, from https://www.symantec.com/content/dam/symantec/docs/reports/istr-21-2016-en.pdf.

Šilić, M., Krolo, J., & Delač, G. (2010). Security vulnerabilities in modern web browser architecture. In *MIPRO, 2010 Proceedings of the 33rd International Convention* (pp. 1240–1245). Red Hook, NY: Curran Associates.

University of Nottingham. (2014). *The Directory of Open Access Repositories – OpenDOAR.* Retrieved June 9, 2016, from http://www.opendoar.org.

Notes

[1] https://www.symantec.com/content/dam/symantec/docs/reports/istr-21-2016-en.pdf.

[2] URL address of the test: https://www.ssllabs.com/ssltest.

[3] URL address of the test: https://www.htbridge.com/ssl.