Data Security and Trustworthiness in Online Public Services: An Assessment of Portuguese Institutions

João Marco C. Silva
INESC TEC &
University of Minho
Portugal
joaomarco@di.uminho.pt

Vítor Fonte
United Nations University (UNU-EGOV) &
University of Minho
Portugal
vff@di.uminho.pt

ABSTRACT

Providing public services through the internet is an effective approach towards an encompassing number of citizens being covered by them and for cost reduction. However, the fast development of this area has fostered discussion and legislation regarding information security and trustworthiness. In addition to security mechanisms for data processed and stored internally, service providers must ensure that data exchanged between their servers and citizens are not intercepted or modified when traversing heterogeneous and uncontrolled networks. Moreover, such institutions should provide means enabling the citizen to verify the authenticity of the services offered. In this way, the present work provides a comprehensive overview regarding the security posture of Portuguese public institutions in their online services. It consists of non-invasive robustness evaluation of the deployed solutions for end-to-end data encryption and the correct use of digital certificates. As a result, we provide some recommendations aiming to enhance the current panorama in the majority of the 111 online services considered in this study.

CCS CONCEPTS

• **Security and privacy → Privacy protections; • Social and professional topics → Privacy policies; • Applied computing** → E-government;

KEYWORDS

Information security, Digital certificates, Privacy, SSL/TLS, Portugal

ACM Reference format:

J. M. C. Silva, V. Fonte. 2019. Data Security and Trustworthiness in Online Public Services: An Assessment of Portuguese Institutions. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance (ICEGOV2019), Melbourne, VIC, Australia, April 3-5, 2019,* 6 pages. https://doi.org/10.1145/3326365.3326411

ICEGOV2019, April 3–5, 2019, Melbourne, VIC, Australia

ACM ISBN 978-1-4503-6644-1/19/04…$15.00
https://doi.org/10.1145/3326365.3326411

1. INTRODUCTION

The continuous growth of global Internet access driven by cheaper personal devices and advances in broadband communication [1] has fostered the use of Information and Communication Technologies (ICTs) as a platform of proximity between public services and citizens. This approach is generically called Electronic Governance (EGOV), and the enhancements leveraged by it include efficiency in deploying a variety of services, a higher number of citizens covered by them, and long-term cost reduction.

Despite the underlying convenience, several of these services require a frequent exchange of personal and sensitive data, for instance, income tax return, medical reports, and social security information. Usually, such data is predominantly transmitted through networks controlled by multiple organizations other than the service provider itself, which might expose essential public services to security incidents.

In response to the security risks involved in providing services over the Internet, some international legislations, such as the European General Data Protection Regulation (GDPR) [2] and similar in Brazil[2], Morocco[3], and Argentina[4] are pushing public and private service providers towards a better control of personal data processed by them.

Some of the most common threats to which users are exposed to when accessing online services are *(i) spoofing* – when an attacker masquerades a trustworthy entity aiming to steal user data; *(ii) information disclosure* – when the confidentiality of users' data is compromised by an attacker eavesdropping the communication channel; and *(iii) data tampering* – when the integrity of data traversing the network is affected by either intentional or accidental modifications [3].

During the last two decades, several technologies have been developed to address these issues by resorting to a cryptography basis to provide secure end-to-end communication, *e.g., Transport Layer Security* (TLS) protocol and online identity certification, *e.g., public key certificates*. Despite their high effectiveness, many services still do not support such technologies or do support

[2] http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm

[3] https://www.cndp.ma/images/lois/Loi-09-08-Fr.pdf

[4] https://www.argentina.gob.ar/sites/default/files/mensaje_ndeg_147-2018_datos_personales.pdf

without following the best practices, for instance, maintaining available legacy versions or broken cryptographic algorithms.

In this context, the main objective of this work is to provide an overview of how Portuguese public institutions address confidentiality, integrity, and trustworthiness in the communication between their online services and the citizens. To pursue this, 111 websites from national institutions were analyzed through non-invasive techniques in order to characterize their cryptographic suites, the existence of known vulnerabilities, and the usage of digital certificates.

The next sections of this paper are organized as follows: Section 2 introduces the main technologies related to end-to-end secure communications and the use of digital certificates as a tool to ensure online identity. Section 3 presents the methodology used to highlight the security mechanisms adopted in communications between public services and citizens. Section 4 discusses the research outcomes, while Section 5 concludes the paper by providing some guidelines about secure communication for online public services.

2. COMMUNICATION SECURITY AND TRUSTWORTHINESS

When offering a service via Internet, a provider has the capacity to implement different mechanisms aiming the security of information processed and stored on its servers. However, due to the open Internet's nature, exchanging data between servers and users implies transmission over networks controlled by different entities, and frequently, with opaque security policies. This scenario rises two main concerns: *(i)* the guarantee that an online service is provided by an authentic entity; and *(ii)* the data transmitted between servers and clients remains private and unchanged while in transit.

The most effective mechanisms designed to cope with both issues have their basis in the field of cryptography. For the first case, the proper use of digital certificates allows high confidence about the authenticity of a service provider, while encrypting the transport-level data flows with robust algorithms is a well-established approach to overcome the second issue. Typically, these approaches are used in conjunction and sustain secure versions of popular protocols, for instance, the Hyper Text Transfer Protocol Secure (HTTPS), which supports the majority of all secure web communication [4]. These technologies are briefly discussed along the next sections.

2.1. Digital certificates

A digital certificate is an electronic document or container file that contains a key value and identifying information about the entity that controls the key [5]. In online communications, the certificate owner (*i.e.,* typically the service provider) sends a copy of its certificate containing a cryptographic public key mathematically related to a private key. By authenticating the embedded key, a user can verify the authenticity of any organization's certificates. Such property is sustained by two main components, namely a *Certification Authority* (CA) and a *Public Key Infrastructure* (PKI).

Certification authorities are entities able to issue a certificate vouching for the identity of any other entity providing online services. It consists in a model of trust relationships, where both subjects in a communication, *i.e.,* the client and the service provider (certificate owner) relies upon the certificate issued by the CA. A recent research [6] has identified 1832 CA worldwide, which are controlled by 683 organizations. Globally, three organizations control 75% of all trusted certificates.

Such certificates are created, managed, distributed, stored and revoked through a *Public Key Infrastructure*, which consists in a set of hardware, software, protocols, legal agreements, processes, and procedures required in secure communications based on public key cryptography. Users rely on such infrastructure in order to validate a certificate received from an online service provider.

Different client-server applications use different types of digital certificates to accomplish their assigned functions. Web servers and web application servers use Secure Sockets Layer (SSL) certificates to authenticate servers via the SSL protocol in order to establish an encrypted SSL session [5]. The SSL protocol is introduced later in this chapter.

It is important to observe that anyone can issue SSL certificates, which are usually called self-signed certificates. However, they will not be trusted automatically by client applications (e.g., web browsers), requiring a CA validation, as it has the ability to issue publicly trusted SSL certificates.

Two popular certificate types are created using Pretty Good Privacy (PGP) [7] and applications that conform to International Telecommunication Union's (ITU-T) X.509 version 3 [8]. Figure 1 presents the structure of a X.509v3 certificate.

2.2. Transport Layer Security (TLS)

Once the service provider's authenticity has been confirmed, a secure communication can be established between the client application and the service server resorting to the same cryptographic principles applied to the certificates. Such secure channel is deployed by using a network protocol designed to prevent eavesdropping, data tampering, and message forgery [9], namely the old SSL, deprecated by the *Transport Layer Security* (TLS), currently in version 1.3 [10].

Some of the main advantages of using such protocols include providing authentication, confidentiality, and integrity in end-to-end communications even in the face of an attacker who has complete control of the network [11], and being application protocol independent, meaning that higher-level protocols (*e.g.,* HTTP) can be deployed on top of the SSL/TLS protocols transparently [12].

The current protocol version is composed of two layers *(i)* a *handshake protocol*, responsible for authenticating entities in a communication, negotiating cryptographic modes and parameters, and establishing shared keying data; and *(ii)* a *record protocol*, that uses the parameters set during the handshake process to protect traffic between the entities [10]. Figure 2 presents a simplified diagram[5] of a secure communication establishment based on the TLS protocol, followed by its description.

[5] Confirmation messages are omitted.

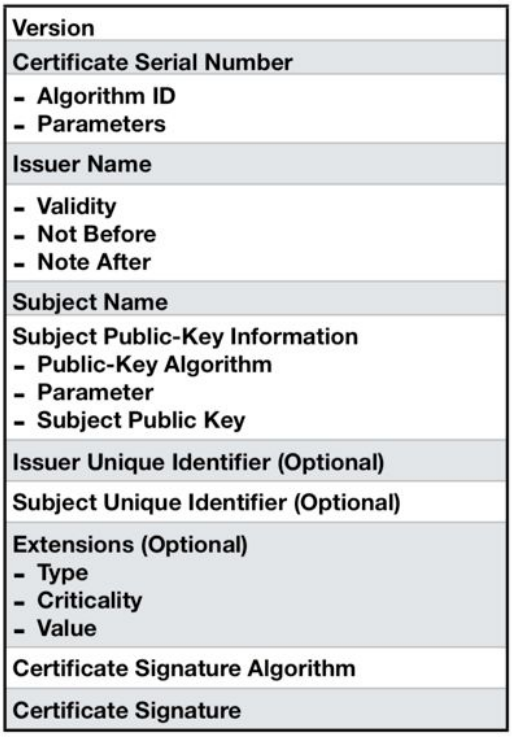

Figure 1: X.509v3 Certificate Structure

1) The communication starts with a "*client hello*" message from the client to the server aiming to establish a connection. This message contains its TLS version, supported cipher suites, algorithm preferences, and a nonce random value *Rc* to be further used;
2) The server responds by sending a "*server hello*" message to the client, along with its choices regarding the client supported cipher suites and algorithms, and its nonce random value *Rs*;
3) The server sends its certificate containing a public key to the client for authentication purposes, as described in Section 2.1. Although not depicted in Figure 2, the server may request a certificate from the client for mutual authentication;
4) The client creates a random Pre-Master Secret and encrypts it with the public key from the server's certificate, sending the encrypted Pre-Master Secret to the server. After receiving the Pre-Master Secret, the server and client each generate the Master Secret and session keys based on the Pre-Master Secret;
5) The client sends "*Change cipher spec*" notification to server to indicate that it will start using the new session keys for hashing and encrypting messages. Based on this message, the server switches its record layer security state to symmetric encryption using the session keys;
6) Client and server can now exchange application data over the secured channel they have established. All messages sent from client to server and from server to client are encrypted using session key.

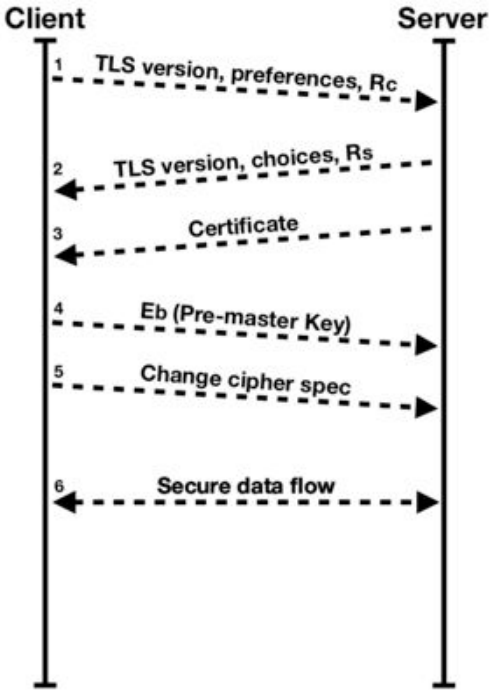

Figure 2: Simplified TLS handshake.

The process of cryptography negotiation represented in Figure 2 is essential towards the security of citizen data when interacting with online public services. Usually, in the presence of a valid certificate and any SSL/TLS protocol version, the client browser shows a padlock indicating that traffic between the browser and server is encrypted. However, it does not mean that the best cryptographic algorithms are being used to secure such communication. Old versions of these protocols still support weak or broken ciphers, for instance, the popular RC4 and MD5, used in symmetric data encryption, and to verify data integrity, respectively, that are no longer recognized as strong choices, due to vulnerabilities to which both are currently exposed [13].

Even considering that most of the main software libraries (e.g., OpenSSL) provide the flexibility to disable a subset of their supported ciphers, it is still common to find online services allowing communications based on weak or broken options. To identify the overall posture of Portuguese public service providers regarding the proper adoption of digital certificates and secure communications is the main objective of this work.

3. METHODOLOGY

Aiming to provide an encompassing overview on how Portuguese public institutions address confidentiality, integrity, and trustworthiness in the communication between their online services and the citizens, a set of non-invasive scanning techniques were applied to 111 domains from 95 public institutions, including municipalities.

The domains analyzed were selected through surveys in online public directories and DNS[6] requests to the specific Second-Level Domain *gov.pt*. Due to security and privacy issues, the domains' name will not be listed in this work, as well as the analysis outcome will be presented aggregately.

The non-invasive scanning techniques adopted consist of analyzing the responses from each domain server during the execution of the SSL/TLS handshake protocol described in Section 2.2. To do so, two main tools were used: *(i) nmap*[7] – a free and open source utility for network discovery and security auditing; and (ii) *SSL Labs*[8] – an online collection of documents and tools developed to assess how TLS/SSL protocols are deployed.

Basically, in step 2 (*i.e.*, Figure 2) the server provides a list of all its supported cipher suites, algorithms, and preferences. It allows assessing whether each server accepts secure connection based on weak or broken algorithms. Such analysis is performed by comparing the resulting list with state of the art in this field and with public vulnerability databases (*e.g.*, NVD[9]).

In step 3 (*i.e.*, Figure 2), the server sends its digital certificate, allowing the analysis of the underlying certification chain in order to ensure the service provider's authenticity.

All the analyzes were performed along August and September of 2018.

4. RESULTS

Following the assessment methodology described in Section 3, the first outcomes show that although being a well-established mechanism to provide secure communications, 20% of the 111 surveyed domains do not support any version of SSL/TLS protocols (see Table 1). It means that, without a cryptographic solution at the application layer, all data being exchanged between clients and servers are unprotected against eavesdropping and tampering, as the underlying traffic traverses the internet in plain text. In addition, they also do not provide any digital certificate, which hampers the user capability of validating the service authenticity.

Considering the remaining 78 services which do support secure communications, 15% do not have a valid digital certificate, either for not providing information about the complete certification chain or for using expired, untrusted or self-assigned certificates. As presented in Table 1, from all 111 services scanned, around 34% do not provide reliable mechanisms validating its authenticity.

Table 1: Summary of the survey.

	Yes	No
Secure transport layer	**80%**	**20%**
Valid certificate	**66%**	**34%**
Weak ciphers	**73%**	**27%**
Presence of vulnerabilities	**65%**	**35%**

Regarding the overall protocol version support, Figure 3 shows that only one service has the most recent and secure TLS protocol (*i.e.*, version 1.3 [10]) implemented. On the other hand, 19 services still support the old and insecure SSLv2 and SSLv3. These protocols versions can be used to attack RSA keys and sites with the same name even if they are on an entirely different server (CVE-2016-0800[10]) and for a man-in-the-middle attacker to obtain clear text data via a padding-oracle attack (CVE-2014-3566), respectively. As detailed in Figure 4, both vulnerabilities represent a high confidentiality risk according to the *Common Vulnerability Scoring System* (CVSS 3.0), meaning that there is a total loss of confidentiality, resulting in all resources within the impacted component being disclosed to the attacker [14].

In addition, 31% of the services also still support legacy TLS v1.0 protocol. Its major vulnerability (*i.e.*, CVE-2011-3389) has been mitigated in modern browsers, however other problems remain [15].

Figure 3 also shows that 28% of the scanned domains support TLS in version 1.1 and 32% support TLS in version 1.2. Although neither has known security issues, only version 1.2 provides modern cryptographic algorithms.

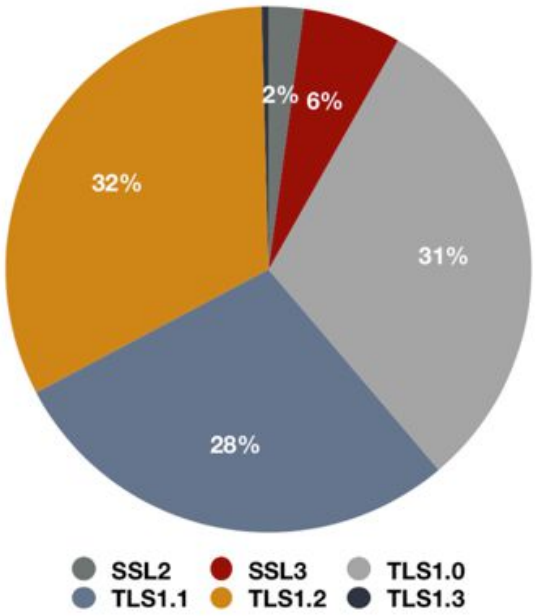

Figure 3: Secure transport protocols

Extending the analysis of confidentiality issues, Figure 3 highlights a significant number of online public services exposed to high impact known vulnerabilities. Namely, 46 are exposed to CVE-2016-6329 or CVE-2016-2183, which allow a remote attacker to obtain clear text data via a birthday attack against a long-duration encrypted session [16].

Another vulnerability that stands out is the CVE-2015-2808, affecting 22 of all scanned services. It is related to a deprecated RC4 algorithm that allows a remote attacker to conduct plaintext-recovery attacks against the initial bytes of a stream by sniffing network traffic [17]. The confidentiality impact in CVSS is "partial", meaning that there is considerable informational

[6] DNS – Domain Name System

[7] https://nmap.org/

[8] https://www.ssllabs.com/

[9] NVD - National Vulnerability Database: https://nvd.nist.gov/

[10] CVE is a list of unique identification numbers and descriptions for publicly known cybersecurity vulnerabilities. Details in https://cve.mitre.org/index.html

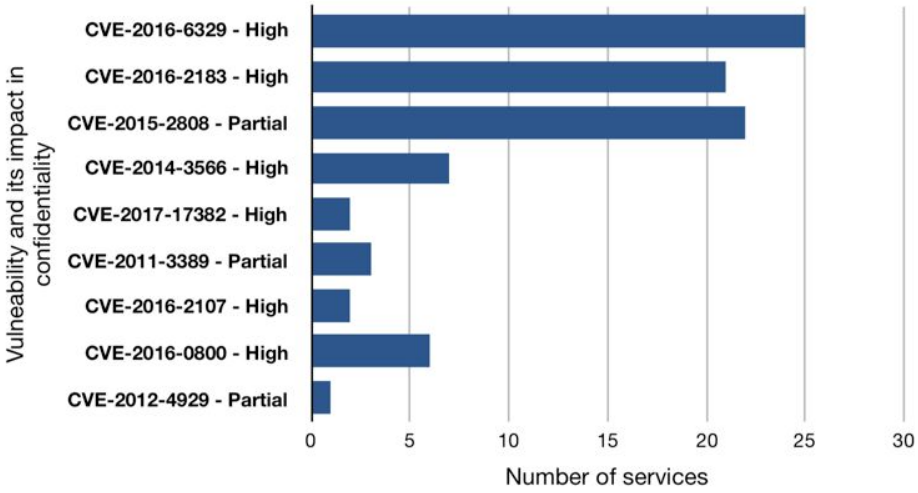

Figure 4: Vulnerabilities and Confidentiality impact.

disclosure, however, the attacker does not have control over what is obtained, or the scope of the loss is constrained [15].

The detailed list of vulnerabilities to which the Portuguese online public services are exposed and their underlying impact on the confidentiality of data being exchanged with users is presented in Figure 4. Overall, 65% of the analyzed services had, at least, one known vulnerability active in their servers by the time of this scanning process.

Beyond the known vulnerabilities, the results have shown that 28 weak ciphers and 7 broken ciphers are still supported in secure communications. In this sense, only 27% of all analyzed domains allow data exchanging based solely in strong ciphers. Table 2 presents the most frequent situations of weak or broken ciphers being supported by different cryptography suites.

Figure 5 summarizes the results of a more comprehensive analysis provided through SSL Lab. It aims at establishing a simple assessment index regarding SSL/TLS server configuration. Such evaluation takes into consideration aspects like certificate trustworthiness, protocol support, key exchange, and cipher strength. The assessment index ranges from *A+* for servers with exceptional configurations, no warnings, and HTTP Strict Transport Security support to *T*, when there is no certificate trust[11].

As presented in Figure 5, 42% of the services considered in the present work have received grade *A+* or *A*, representing public services with solid posture regarding the security of citizen data being exchanged through public networks. However, it is not observed for the majority of the services currently offered to Portuguese citizens, as manifold sensitive problems are observed for the remaining 58% of analyzed servers.

5. DISCUSSION AND CONCLUSIONS

The fast growth of public services being offered through the internet has also sparked concerns regarding privacy and security of data traversing public and uncontrolled networks. Moreover, providing mechanisms for authenticity verification is a key aspect towards citizen engagement to such services.

By resorting to non-invasive scanning methodologies, this work has identified that most of the main online public services provided in Portugal have serious problems regarding data security and trustworthiness.

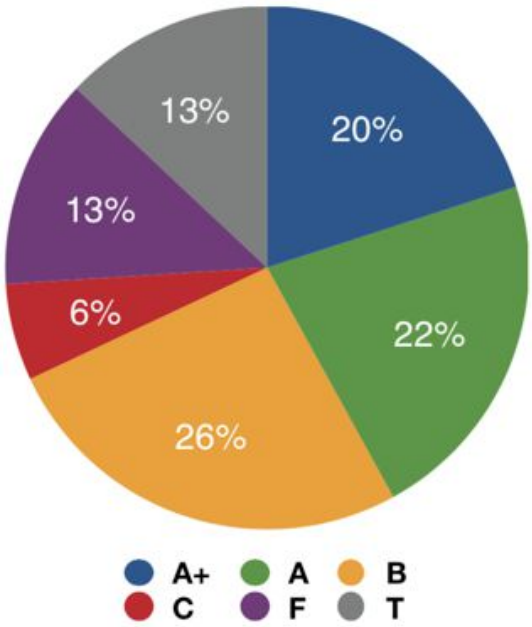

Figure 5: Global result from SSL Lab.

One of the main problems found is related to the support of legacy protocols and, consequently, weak or broken ciphers. Considering that most of the softwares providing secure

[11] More details regarding the rating criteria can be find in https://github.com/ssllabs/research/wiki/SSL-Server-Rating-Guide

communications to online services are of easy and flexible configuration, even after the service deployment, such scenario might indicate a lack of continuous evaluation of the risks related to the cryptographic suites selected in the service design stages.

Table 2: Availability of weak or broken ciphers.

Cipher	Status	Number of suites supporting it
TLS_RSA_WITH_AES_128_CBC_SHA	Weak	58
TLS_RSA_WITH_AES_256_CBC_SHA	Weak	57
TLS_RSA_WITH_AES_256_CBC_SHA256	Weak	47
TLS_RSA_WITH_AES_128_CBC_SHA256	Weak	46
TLS_RSA_WITH_AES_128_GCM_SHA256	Weak	39
TLS_RSA_WITH_AES_256_GCM_SHA384	Weak	38
TLS_RSA_WITH_AES_256_CCM_8	Weak	32
TLS_RSA_WITH_RC4_128_SHA	Broken	20
TLS_RSA_WITH_RC4_128_MD5	Broken	19
TLS_ECDHE_RSA_WITH_RC4_128_SHA	Broken	6

Following reports from public vulnerability databases or relying on a centralized group of experts able to handle computer security incidents (*i.e.*, a Computer Emergency Response Team – CERT) could provide relevant inputs to professionals in charge of the diverse online services across the country at a low cost and high efficiency. It also should involve investing in continuous people qualification [18], as this area is in constant and fast development.

In addition, regular auditing processes have the potential to reduce the service exposition to known vulnerabilities. It might be conducted through open source and free tools, such *nmap* and *SSL Lab* used in the present study. Furthermore, a deeper perspective of potential security issues might be achieved through penetration testing activities.

Another sensitive issue found in this research is the lack of or the use of untrusted digital certificates, which hampers the capacity of verifying whether an online service is provided by an authentic institution. This aspect raises serious concerns regarding attacks deployed to steal sensitive data from users.

Overall, these finds show the same issues pointed out by previous related work [6], which observed that, worldwide, half of trusted leaf certificates contain an inadequately secure 1024-bit RSA key in their trust chain and that CAs were continuing to sign certificates using MD5 as late as April 2013.

As future work, we intend to contact each institution considered in the present study with the aim at providing a detailed report comprehending the issues identified on their security posture regarding data being exchanged between their servers and citizens.

ACKNOWLEDGEMENTS

This paper is a result of the project SmartEGOV: Harnessing EGOV for Smart Governance (Foundations, Methods, Tools) NORTE-01-0145-FEDER-000037, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (EFDR).

REFERENCES

[1] OECD. 2018. Internet access (indicator). Technical Report. https://doi.org/10.1787/69c2b997-en

[2] 2016. Regulation (EU) 2016/679of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da. Official Journal of the European Union L119 (may 2016), 1–88.

[3] Tim Dierks and Eric Rescorla. 2008. The transport layer security (TLS) protocol version 1.2 - RFC 5246. Technical Report. https://doi.org/10.17487/RFC5246

[4] Benjamin Vander Sloot, Johanna Amann, Matthew Bernhard, Zakir Durumeric, Michael Bailey, and J Alex Halderman. 2016. Towards a Complete View of the Certificate Ecosystem. In Proceedings of the 2016 Internet Measurement Conference (IMC '16). ACM, New York, NY, USA, 543–549. https://doi.org/10.1145/2987443. 2987462

[5] Michael E. Whitman and Herbert J. Mattord.2011. Principles of information security. Cengage Learning.

[6] Zakir Durumeric, James Kasten, Michael Bailey, and J Alex Halderman. 2013. Analysis of the HTTPS certificate ecosystem. In Proceedings of the 2013 conference on Internet measurement conference. ACM, 291–304.

[7] Jon Callas, Lutz Donnerhacke, Hal Finney, David Shaw,and Rodney Thayer.2007. OpenPGP message format. Technical Report.

[8] Stefan Santesson, Magnus Nystrom, and Tim Polk. 2004. Internet x. 509 public key infrastructure: Qualified certificates profile (RFC 3739 IETF). Technical Report.

[9] Tim Dierks and Eric Rescorla. 2008. The transport layer security (TLS) protocol version 1.2 - RFC 5246. Technical Report. https://doi.org/10.17487/RFC5246

[10] Eric Rescorla. 2018. The transport layer security (TLS) protocol version 1.3 – RFC 8446. Technical Report. RFC - Proposed Standard (IETF Stream).

[11] Eric Rescorla and Brian Korver. 2003. Guidelines for writing RFC text on security considerations - RFC 3552. Technical Report. RFC - Proposed Standard (IETF Stream).

[12] Andrew S Tanenbaum and David J Wetherall. 2010. Computer Networks (5th ed.). Prentice Hall Press, Upper Saddle River, NJ, USA.

[13] Daniel A Menascé. 2003. Security performance. IEEE Internet Computing 7, 3 (2003), 84–87.

[14] C S Team. 2015. Common Vulnerability Scoring System v3.0: Specification Document. First. org (2015).

[15] Peter Mell, Karen Scarfone, and Sasha Romanosky. 2007. A complete guide to the common vulnerability scoring system version 2.0. In Published by FIRST-Forum of Incident Response and Security Teams, Vol. 1. 23.

[16] Paul Kirchner. 2011. Improved Generalized Birthday Attack. IACR Cryptology ePrint Archive 2011 (2011), 377.