Authenticity and versioning of learning objects using the digital signature infrastructure of Costa Rica

Isaac Alpizar-Chacon*, Mario Chacon-Rivas†
TEC Digital
Instituto Tecnologico de Costa Rica
Cartago, Costa Rica
*Email: {*ialpizar,†machacon}@tec.ac.cr*

***Abstract*—There are a large number of educational resources available on the Internet through different repositories. However, authenticity methods of learning objects has not been widely adopted by repositories, making it difficult to verify the authors of the content. To solve this problem, our goal is *to aid users and authors of learning objects to have a trusted verification of authenticity and versioning of learning objects by applying digital signatures into the lifecycle of learning objects*. In our propose, after a learning object is uploaded to a repository, the author presents his/her digital certificate, a special LOM record is created with the information from the author's certificate, and then it is signed to prove its authenticity. The signature is produced using digital certificates issued by the digital signature infrastructure of Costa Rica, which guarantees legal binding and trust over time. We hope to propose a generic framework than can be implemented in any repository.**

***Keywords*—Learning Objects, Digital Signatures, Authenticity, Learning Object Repositories**

1. Introduction

Despite the large number of educational resources available in Internet, the concept of authenticity of content in learning objects has not been widely implemented by systems in order to enforce the authorship and integrity of the information [1]. The MERLOT[1] repository contains over 40,000 materials in 19 different categories, ARIADNE[2] aggregates around 830,000 learning object metadata elements and LAFLOR[3] exposes more than 50,000 educational materials, just to give some examples. When users search learning objects (LO) in those repositories they can get hundreds or even thousands or results. Those results often do not show validated information regarding the author and provenance of the learning object. Even many repositories allow users to submit their own material after a simple registration, and although this practice helps to share and reuse knowledge, quality of the content is not guarantee and is very difficult to trust the authenticity of the author who uploads the learning object.

The goal of this research is *to aid users and authors of learning objects to have a trusted verification of authenticity and versioning of learning objects by applying digital signatures into the lifecycle of learning objects*.

Our proposed approach uses a trusted repository where authors can submit their learning objects. The author uploads and introduces all the information regarding the learning object, and then the repository requests the digital certificate from the user. New data elements are added to the LOM record of the learning object, these elements contain the extracted credentials of the author from the digital certificate. Finally, the author signs the XML LOM registry using his/her digital certificate. The repository stores the learning object with the signed information. When a user accesses the learning object, the repository validates the signature, and displays an authenticity check.

We take advantage of the national digital signature infrastructure to propose this new service to the education community of Costa Rica. The government makes available to citizens and residents access to digital certificates through Registration Offices throughout the country at a low cost. Our proposal relies in this infrastructure for the generation, management, and distribution of the necessary digital certificates. Digital signature and electronic documents in Costa Rica have the same legal status and probative value as handwritten signatures and electronic documents [2], which gives confidence in our authenticity model.

Although this research is a work in progress, we hope to propose a generic framework than can be implemented in any repository to give confidence to the users by providing authenticity of learning objects. Many other countries like Mexico[4], Brazil[5], Spain[6], and Belgium[7] have national digital signature infrastructures and could implement the framework described in this paper.

In Section 2 we present related work that has been done on the subject. Section 3 presents and explains all the details

1. https://www.merlot.org
2. http://ariadne.grnet.gr/
3. http://laflor.laclo.org/
4. http://www.firmadigital.gob.mx/
5. http://www.iti.gov.br/icp-brasil
6. http://www.dnielectronico.es/
7. http://eid.belgium.be/en

of the proposed approach. The limitations are discussed in Section 4. Finally, in Section 5 we state the main conclusions of the research and mention future work to be conducted.

2. Background and Related Work

Authenticity of learning objects using digital certificates has been explored before. In [3] the authors proposed to use digital certificates to sign learning objects that are represented using Sharable Content Object Reference Model (SCORM) inside a Learning Content Management System (LCMS). Different security models are explored in the paper and a new model is proposed. The model adds an XML Signature and XML Encryption element to the SCORM manifest file and inside the Rigths element of LOM. The LCMS is in charge of the generation, management and distribution of the digital certificates. The proposed approach is later implemented in Moodle in [1]. This second paper focuses on how the management of digital certificates must be carried out in Moodle and how the LCMS should act as certificate authority.

Our approach also uses digital certificates to sign learning objects, but differs in several aspects. First, we use and trust the digital signature infrastructure of Costa Rica, which is accessible to all citizens, residents, institutions and organizations throughout the country, and not only to specific users of a LCMS. Second, we apply the digital signature directly to a complete LOM xml file, which is a more general standard than SCORM. Third, we use the XAdES format to encode the signed file, which extends XML Signature to making it suitable for Advanced Electronic Signatures (see section 3.1.1). Finally, our approach can be applied independently by the author (without a LCMS) or inside a Learning Object Repository or Institutional Repository (see section 3.1.3).

Versioning of learning objects has been describe in [4], [5], [6], [7]. Typically versioning of a learning object is done in both the content and the metadata of the resource. This kind of versioning allows to compare two versions of a learning object, and obtain all the changes made from one version to the other. In contrast, we propose a simpler versioning where we only identify, authenticate and trace the different versions of a learning object. It is up to the user the identification of the exact changes that were made in the different versions of a learning object.

In this paper a LOM application profile is proposed. Application profiles are "meta-data element sets that are either abbreviated versions of complete standards or are a heterogeneous mix of elements drawn from different metadata schemata" [8]. Application profiles are developed to accomplish requirements that are specific to an application, within a community. Some examples of LOM application profiles are:

1) LOM-ES: developed for the education sector in Spain [9],
2) ANZ-LOM: developed for the education sector in Australia and New Zealand [10],
3) FAO Learning Object Resources Metadata Application Profile: developed to describe agricultural learning resources [11], and
4) SG-LOM: an IEEE LOM application profile to describe serious games [12].

Finally, some authors have proposed methods to prove the ownership and to protect learning objects from unauthorized access [13], [14]. Those kinds of restrictions are outside the scope of this work. If the enviroment requires protection of learning objects, we trust that our approach is compatible with such techniques.

3. Proposed Approach

As mention above, we propose a framework to check the authenticity of learning objects using digital certificates, specifically the digital signature infrastructure of Costa Rica. Additionally, the framework allows to verify the authors who have contributed in different versions of a learning object through time. Verification of authenticity can be done by a trusted learning object repository that can automatically check the signatures on the metadata of learning objects or by the users themselves if they have the necessary means (software, certificates, knowledge, etc.).

In this section we will first present the different aspects that are part of the framework and then finish with the proposed signature and validation process.

3.1. Key Components

3.1.1. Advanced Electronic Signatures. An electronic signature as defined in the European Directive [15] is "data in electronic form which are attached to or logically associated with other electronic data and which serve as a method of authentication". The Directive also defines the advanced electronic signature as an extended electronic signature that meets special requirements to accomplish authentication (identity of the author is verificable), integrity (the information has not been modified), and non-repudiation (it is not possibly to deny the authenticity of the applied signature). Readers are referred to [16] for more information on these concepts.

There are three types of Advanced Electronic Signatures:

1) PAdEs [17] that is used to sign PDF files,
2) CAdES [18] that is used to sign any binary data, and
3) XAdES [19] that is used to sign XML files.

XAdEs is the format used in this approach because LOM records are usually created in XML [20] and because many repositories (ARIADNE[8], LA FLOR[9], DSpace[10], EPrints[11], etc.) expose metadata using the OAI-PMH protocol [21] which provides an XML response.

8. http://ariadne.grnet.gr/ariadne-repository/services/oai
9. http://laflor.laclo.org/admin/services/oai
10. https://wiki.duraspace.org/display/DSDOC5x/OAI
11. http://wiki.eprints.org/w/OAI

3.1.2. Costa Rica Public Key Infrastructure. Since 2005 Costa Rica has the Law on Certificates, Digital Signature and Electronic Documents - Law 8454 [2], which provides the jurisdiction for the issuance and use of digital signature certificates in the country. The national infrastructure involves the Ministry of Science, Technology and Telecommunications (MICITT), and the Central Bank of Costa Rica which give great support and confidence to the national Public Key Infrastructure. Costa Rica citizens and residents can obtain a smart card[12] containing the digital certificate in around sixteen different Registration Offices across the country.

Updated statistics[13] from the first quarter of 2016 show that 153,123 persons have obtained a smart card with their digital certificate since 2009. This gave us the necessary confidence to rely in this infrastructure to provide a framework for the authenticity and versioning of learning objects in the country.

An important aspect that affects and makes more robust the proposed approach is that the official digital signature format for XML documents in Costa Rica is XAdES-X-L [19]. This format includes extended validation data to provide more confident long term signatures.

3.1.3. Learning Object Repositories and Institutional Repositories. Learning Object Repositories (LOR) [22] are digital libraries to store, manage and share digital educational resources. Institutional Repositories (IR) [23] are also digital libraries to store scholarly material from a particular institution. Collections inside a IR can be dedicated to learning objects and other educational resources.

The proposed approach uses a repository (LOR or IR) as the main tool to perform the signature and the validation of LO's. Inside a repository when an author is submitting a LO, there are two possible ways: 1) self-submission (authors submit their own work without a third person), and 2) traditional submission (authors give their work to a third person, who is in charge of submission). We focus on self-submission.

It is possible to adapt the proposed process in two aspects: first, so the author can sign the object by himself (without a repository in the middle) and then share it at his discretion, and second to use traditional submission as the means to upload the LO to the repository.

3.1.4. LOM Record. The actual digital signature is produced on the XML file containing the LOM record of the LO instead of the bitstreams of the LO. This is because the LOM record contains elements that tides the author to the learning object. It is necessary to add information to LOM in order to accomplish the authenticity and versioning of the learning object. To this purpose a LOM application profile, called DS-LOM, was created. The DS-LOM application profile maintains all the elements of LOM, but adds new obligatory elements and extends some value spaces. Table 1 shows the fields added to LOM in DS-LOM and explains the purpose of each element. Table 2 shows fields that are part of LOM but should be used in a specific way. One of the most important added elements is 2.3.4 Hash. The hash field stores the SHA-1 hash of all the bitstreams that are part of the LO, so the author can be linked to specific files. Other elements allow to identify previous versions of the learning object (7.2.3 version), the country in which the signature is legally valid (4.8.1 jurisdiction), and the repositories capable to validate the learning object (4.8.2 repository).

3.1.5. vCard Format. LOM uses the vCard 3.0 format[14] to encode the information about the entities that contribute to the creation of the LO. We need to identify the author of the LO with the same information that is present in his/her digital certificate. For that purpose, we defined a template to encode the data in the digital certificate into the structure of vCard 3.0. An example of this structure is presented in figure 1. The most important field in the digital certificate is 'Serial Number' because it uniquely identifies the author legally in the country. The serial number of the author is encoded in the vCard using the NICKNAME property due to the lack of a specific property to store the ID of a person. The given name, surname, country, organization and organizational unit of the author are encoded in the vCard as well.

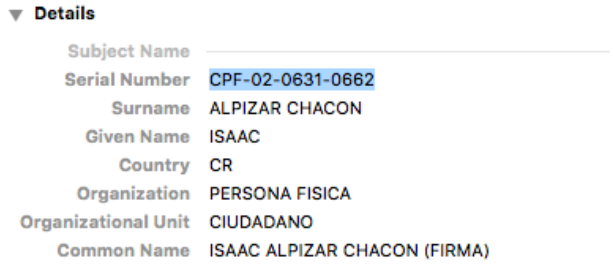
3.1.6. Java Applet. After the information about the LO has been introduced in the repository, the LOM record is generated in the server-side, and the author needs to sign it from the client-side. To accomplish this scenario, the author needs to be able to sign documents over a web browser. Mozilla proposed and used for some time the JavaScript Crypto library[15], but it is deprecated since Firefox 34 (end of 2014). Microsoft developed the CAPICOM ActiveX control, but is discontinued since 2009[16] and works only with Internet Explorer. The W3C has worked in a standard for the industry, called the Web Cryptography API[17], but it does not have support for smart cards, it is still a draft and is not yet implemented in all mayor web browsers[18].

Accordingly to these scenarios, we propose to use a Java Applet to sign the LO. A Java Applet can access smart cards and be used to sign any document [24], [25]. Java is supported in many browsers (unsupported in Chrome from version 45) and although is not a standard solution it is a viable implementation technology. Other technologies like Java Web Start[19] could help to overpass Chrome's restrictions and execute the applet outside the browser. This option will be explored in the future.

12. http://www.smartcardalliance.org/smart-cards-faq/#smartcard
13. http://www.bccr.fi.cr/firma_digital/estadisticas.html
14. https://tools.ietf.org/html/rfc2425, https://tools.ietf.org/html/rfc2426
15. https://developer.mozilla.org/enUS/docs/Archive/Mozilla/JavaScript_crypto
16. https://blogs.msdn.microsoft.com/karinm/2009/01/18/capicom-dll-removed-from-windows-sdk-for-windows-7/
17. https://dvcs.w3.org/hg/webcrypto-api/raw-file/tip/spec/Overview.html
18. https://www.chromestatus.com/feature/5030265697075200
19. http://www.oracle.com/technetwork/java/javase/javawebstart/index.html

```
BEGIN:VCARD
VERSION:3.0
N:ALPIZAR CHACON;ISAAC;;;
FN:ISAAC ALPIZAR CHACON (FIRMA)
NICKNAME:CPF-02-0631-0662
ORG:PERSONA FISICA;CIUDADANO
ADR;TYPE=HOME:;;;;;;CR
NOTE:/serialNumber=CPF-02-0631-0662/
SN=ALPIZAR CHACON/GN=ISAAC/C=CR/
O=PERSONA FISICA/OU=CIUDADANO/
CN=ISAAC ALPIZAR CHACON (FIRMA)
END:VCARD
```

(a) Example of proposed vCard format



(b) Example of digital certificate information

Figure 1: Author information in digital certificate and vCard format

3.2. Signature Process

Here we present the proposed self-submission flow in a repository to sign learning objects.

3.2.1. New LO. When an author is submitting a new LO the process is as follows:

1) Author logs in into the LOR or IR using any available authentication method.
2) Author selects the type of submission, in this case: new LO.
3) Author fills all elements in the submission form, except the common "author" field. This information will be extracted from the author's digital certificate.
4) After the author has given all the relevant information about the LO (this will vary between repositories), the system asks the user to digital sign the learning object. The signature is performed using a signing Java Applet (see 3.1.6). The applet obtains the author's digital certificate, extracts the identification information, completes the DS-LOM record of the LO and then the author signs it.
5) The repository takes the signed XML document, adds some information to make it a XAdES-X-L document and then stores it.
6) Finally, the author reviews the signed information and the process ends.

3.2.2. New version of existing LO. When the author is submitting a new version of an existing learning object, two steps change respect to 3.2.1:

2) Author selects: new version of LO.
3) Author must fill all elements as a new LO, plus additional information. This information identifies the latest version of the learning object on which the new version was built. If the repository contains all the versions of the learning object, author only needs to provide the identifier of the latest version. Otherwise, the author need to fill the identifier, version number and the hash of the LOM record of the learning object. The author should also upload all previous (including the original) versions of the learning object.

3.3. Validation Process

When a user access a learning object that has a digital signature, the repository needs to validate the signature. If the signature is correct, the repository shows a check for the authenticity of the learning object along with the author's information.

If the learning object is a new version and the signature is valid, the repository needs to validate the chain of all previous versions of the learning object. For each learning object in the chain, the repository shows a link to the resource and the corresponding author's information.

4. Limitations

This work has the following limitations:

1) The approach relies in the existing public key infrastructure of Costa Rica. To implement the proposed approach in a different country, it needs to have a similar infrastructure. If such requirement does not exist, it is possible to build an entity to create, manage and distribute digital certificates. The creation of that body will consume significant money and time.
2) If the growth of persons with digital certificate slows, it is possible that many authors can not sign their own learning objects.
3) We trust the author when he/she claims the authorship and ownership of a learning object with his/her digital signature. It may happen that a dishonest author signs a learning object that does not belong to him/her. Plagiarism detection and the enforcement of intellectual property are beyond the scope of this paper.

TABLE 1: New fields that have been added in DS-LOM

Nr	Name	Explanation	Size	Value Space	Data Type
2.3.1	Role	Two more options were added to the value space: Original author and Contributing author. This element identifies if the learning object is new, or it is a new version of an existing learning object.	1	Value space from LOMv1.0 plus: 'original author' and 'contributing author'	Vocabulary (State)
2.3.4	Hash	Base64 SHA-1 hash of the bitstream or bitstreams that represent the content of the learning object. If there are two or more bitstreams the hash is calculated over the zip file containing all bitstreams.	1	Repertoire of ISO/IEC 10646-1:2000	CharacterString (min: 28 char, max: 28 char)
7.2.3	Version	Version of the previous learning object. This element is the value of 2.1 of the previous version of the LO. This field is obligatory for new versions of a LO.	1	-	LangString (smallest permitted maximum: 50 char)
7.2.4	Hash	Base64 SHA-1 hash of the bitstream or bitstreams that represent the content of the previous learning object. This field is the value of 2.3.4 of the previous LO. This field is obligatory for new versions of a LO.	1	Repertoire of ISO/IEC 10646-1:2000	CharacterString (min: 28 char, max: 28 char)
4.8	Validation	Aggregate data element to indicate where the validation of the digital signature should be done.	1	-	-
4.8.1	Jurisdiction	Country in which the applied digital signature is legally valid nationwide. This information helps to locate the necessary tools and CA certificates to validate the signature manually.	smallest permitted maximum: 100 items.	Country code from the code set ISO 3166-1:1997. Two-letter country code in upper case	CharacterString (min: 2 char, max: 2 char)
4.8.2	Repository	Trusted repository capable to validate the applied digital signature.	smallest permitted maximum: 5 items.	Repertoire of ISO/IEC 10646-1:2000	CharacterString (smallest permitted maximum: 1000 char)

TABLE 2: Fields with specific behaviour in DS-LOM

Nr	Name	Explanation	Size	Value Space	Data Type
2.3.2	Entity	The vCard information about the author should follow a defined structured that represents the personal information in the digital certificate of the author. There should be a data element for each of the 'orignal author' or 'contributing author'.	smallest permitted maximum: 40 items	vCard, as defined by IMC vCard 3.0 (RFC 2425, RFC 2426) and as defined in 3.1.5	CharacterString (smallest permitted maximum: 1000 char)
7.1	Kind	When describing a new version of a learning object, this element is obligatory and should use specific value space.	1	isversionof, isbasedon	Vocabulary (State)

5. Conclusions and Future Work

We have proposed a framework to validate the authenticity and versioning of a learning object using digital signatures. The approach includes the author's information in a extended LOM record (DG-LOM application profile), the hash of the contents of the learning object, and the id of previous versions of the resource if any. The XML registry is signed by the author to prove the authenticity of the learning object. The signed learning object may reside inside a trusted repository in order to facilitate future validations of the digital resource. Although the approach was develop using the digital signature infrastructure of Costa Rica, it can be applied in any other context where a Public Key Infrastructure is available.

In order to attack the lack of systems that enforce and implement authenticity of learning objects, the next step of this work is to implement the proposed approach in RepositorioTEC[20] (IR of Instituto Tecnologico de Costa Rica[21], a public university specialized in engineering and science), which is based on DSpace. We want to create a module that can be used by other instances of DSpace. After the development of the module, we would like to validate if the LO's using this framework obtain more trust in their authenticity among users compared to LO's without proven authenticity. Finally, we expect to take this approach further and propose a trusted federation of repositories than can

20. http://repositoriotec.tec.ac.cr/
21. http://www.tec.ac.cr/Paginas/index.html

aggregate different national repositories from countries with digital signature infrastructures.

References

[1] P. A. Gaona, C. E. M. Marín, and H. W. Gonzalez, "Hacia una propuesta de mecanismos para la autenticidad de objetos de aprendizaje en plataformas lcms," *Ingeniería*, vol. 19, no. 1, 2014.

[2] "Ley de certificados, firmas digitales y documentos electrnicos [the law on certificates, digital signature and electronic documents]," Ley 8454 [Law 8454], LA ASAMBLEA LEGISLATIVA DE LA REPBLICA DE COSTA RICA [Legislative Assembly of Costa Rica], oct 2005. [Online]. Available: http://www.firmadigital.go.cr/Documentos/ley%208454.pdf

[3] C. E. M. Marín, E. E. G. García, and P. A. G. García, "Plataforma de seguridad basado en autenticidad de contenidos sobre conjunto de especificaciones scorm," *Revista Ingeniería y Competitividad*, vol. 12, no. 2, pp. 51–68, jun 2011. [Online]. Available: http://revistaingenieria.univalle.edu.co:8000/index.php/inycompe/article/view/179

[4] C. Brooks, J. Cooke, and J. Vassileva, "Versioning of learning objects," in *Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference on*, July 2003, pp. 296–297.

[5] A. S. Krebs, "Delta versioning for learning objects," U.S. Patent US8 121 985 B2.

[6] R. Green, S. Basu, S. Shankar, K. Rajamani, H. Hung, and J. Lee, "System and method for managing versions of metadata," U.S. Patent US8 005 792 B2.

[7] W. Theilmann, M. Altenhofen, and W. Gerteis, "Versioning electronic learning objects," U.S. Patent US20 040 126 750 A1.

[8] IMS Global Learning Consortium. (2006, aug) Ims metadata best practice guide for ieee 1484.12.1-2002 standard for learning object metadata. IMS Global Learning Consortium. [Online]. Available: https://www.imsglobal.org/metadata/mdv1p3/imsmd_bestv1p3.html#1634635

[9] M. Canabal, A. Sarasa, and J. C. Sacristán, "Lom-es: Un perfil de aplicación de lom," in *Simposio SPEDECE, V Simposio Pluridisciplinar sobre Diseño y Evaluación de Contenidos Educativos Reutilizables, Salamanca*, 2008.

[10] Education Services Australia. (2015, mar) Anz-lom metadata application profile. Education Services Australia. [Online]. Available: http://www.ndlrn.edu.au/verve/_resources/ANZ_LOM_map.pdf

[11] H. Stuempel, G. Salokhe, A. Aubert, J. Keizer, A. Nadeau, S. Katz, and S. Rudgard, *Metadata and Semantics*. Boston, MA: Springer US, 2009, ch. Metadata Application Profile for Agricultural Learning Resources, pp. 499–507. [Online]. Available: http://dx.doi.org/10.1007/978-0-387-77745-0_49

[12] Y. E. Borji and M. Khaldi, "An ieee lom application profile to describe serious games 'sg-lom'," *International Journal of Computer Applications*, vol. 86, no. 13, pp. 1–8, January 2014.

[13] O. A. Santos and F. M. S. Ramos, "Proposal of a framework for internet based licensing of learning objects," *Comput. Educ.*, vol. 42, no. 3, pp. 227–242, Apr. 2004. [Online]. Available: http://dx.doi.org/10.1016/j.compedu.2003.07.002

[14] H. Madhour, M. A. Sfaxi, M. W. Forte, and S. G. Helie, "Ownership detection and protection for learning objects," in *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, July 2006, pp. 784–788.

[15] C. o. t. E. U. European Parliament, "Directive 1999/93/ec of the european parliament and of the council of 13 december 1999 on a community framework for electronic signatures," *Official Journal of the European Communities. Legislation*, vol. 43, pp. 12–20, jan 2000.

[16] C. Adams and S. Lloyd, *Understanding PKI: Concepts, Standards, and Deployment Considerations*, 2nd ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2002.

[17] *Electronic Signatures and Infrastructures (ESI); PDF Advanced Electronic Signature Profiles; Part 1: PAdES Overview - a framework document for PAdES*, European Telecommunications Standards Institute Std. 102 778-1.

[18] *Electronic Signatures and Infrastructures (ESI); CMS Advanced Electronic Signatures (CAdES)*, European Telecommunications Standards Institute Std. 101 733.

[19] *XML Advanced Electronic Signatures (XAdES)*, European Telecommunications Standards Institute Std. 101 903.

[20] *IEEE Standard for Learning Technology-Extensible Markup Language (XML) Schema Definition Language Binding for Learning Object Metadata*, IEEE Std. 1484.12.3, 2005.

[21] C. Lagoze and other. (2015, jan) The open archives initiative protocol for metadata harvesting. [Online]. Available: https://www.openarchives.org/OAI/openarchivesprotocol.html

[22] R. Lehman, "Learning object repositories," *New directions for adult and continuing education*, vol. 2007, no. 113, pp. 57–66, 2007.

[23] M. Ware, "Institutional repositories and scholarly publishing," *Learned publishing*, vol. 17, no. 2, pp. 115–124, 2004.

[24] S. Nakov. (2006, feb) Java applet for signing with a smart card. Developer.com Network. [Online]. Available: http://www.developer.com/java/other/article.php/3587361/Java-Applet-for-Signing-with-a-Smart-Card.htm

[25] Z. Chen, *Java Card Technology for Smart Cards: Architecture and Programmer's Guide*, ser. Addison-Wesley Java Series. Addison-Wesley, 2000.