TRENCADIS – Secure Architecture to Share and Manage DICOM Objects in a Ontological Framework Based on OGSA

Ignacio BLANQUER[1], Vicente HERNANDEZ[1], Damià SEGRELLES[1], Erik TORRES[1]

[1]Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas, Universidad Politécnica de Valencia, Camino de Vera S/N, 46022 Valencia Spain (phone: +34 96 387 7007 ext.88254; Fax: +34 96 387 7274; e-mails: iblanque@dsic.upv.es; vhernand@dsic.upv.es; dquilis@itaca.upv.es; etorres@itaca.upv.es)

Abstract. Today most European healthcare centers use the digital format for their databases of images. TRENCADIS is a software architecture comprising a set of services as a solution for interconnecting, managing and sharing selected parts of medical DICOM data for the development of training and decision support tools. The organization of the distributed information in virtual repositories is based on semantic criteria. Different groups of researchers could organize themselves to propose a Virtual Organization (VO). These VOs will be interested in specific target areas, and will share information concerning each area. Although the private part of the information to be shared will be removed, special considerations will be taken into account to avoid the access by non-authorized users. This paper describes the security model implemented as part of TRENCADIS. The paper is organized as follows. First introduces the problem and presents our motivations. Section 1 defines the objectives. Section 2 presents an overview of the existing proposals per objective. Section 3 outlines the overall architecture. Section 4 describes how TRENCADIS is architected to realize the security goals discussed in the previous sections. The different security services and components of the infrastructure are briefly explained, as well as the exposed interfaces. Finally, Section 5 concludes and gives some remarks on our future work.

Keywords. Grid, Security, DICOM

Introduction

The use of digital medical images in hospital environments has changed the way in which radiologists work and cooperate. The generalization of Digital Imaging and Communications in Medicine [1] (DICOM), as a world-wide standard for the transmission and exchange of medical images, has made it possible to share images across a wide set of users and applications. Today most European healthcare centers use the digital format for their databases of images (PACS, Hard Disk Shared Directories, etc…). These databases can be located at different physical sites into a single medical centre or even also in different centers. Therefore, this diversity in the source of images, the distributed location of the storages and managing systems difficulties the transparent sharing of images and the development of secure collaborative environments, where new applications of Medical Imaging (Advanced Image-based Diagnosis, Non-Affine Registration/Fusion Applications, Context-Based searching of DICOM images etc...) can be built.

On the other hand, DICOM does not only imply images coming from the radiology departments, other type of images and movies (dermatology, endoscopes, etc.) and other type of information such as radiology reports are being coded into DICOM. DICOM Structured Reporting [2] (DICOM-SR) codes and integrates radiology reports with seamless references to findings and Regions of Interests on the associated images. Structuring radiology reports offers a comparable way to code reports enhancing the capability of tools to search and to extract knowledge.

To tackle this challenge, a software architecture, namely TRENCADIS[1] [3] [4] was developed and implemented, comprising also a set of services as a solution for interconnecting, managing and sharing selected parts of medical DICOM data for the development of training and decision support tools. The organization of the distributed information in virtual repositories is based on semantic criteria. Different groups of researchers could organize themselves to propose a Virtual Organization (VO). These VOs will be interested in specific target areas (e.g. pediatric oncology), and will share information (studies and reports) concerning each area. Subsets of those images could be obtained for a specific study (e.g. neuroblastoma). Finally, in each subset, users can make complex queries (e.g. male patients above one year with irregular findings of more than 3 mm). Of course, the sharing of the information must be secure and within the VO. Although the private part of the information to be shared will be removed, special considerations will be taken into account to avoid the access by non-authorized users (even those with administrative privileges at remote sites). It is important to notice that data security is a key requirement for biomedical grid applications, since not only the obvious technical requisite of ensuring integrity and validity of computations must be guaranteed, but also the necessity of being liable to heterogeneous national legal regulations and developing procedures to be accepted by the medical community [5]. Once the architectural design [6], the ontological framework [4] and the management of DICOM SR [3] have been completed, this paper describes the security model implemented as part of the architecture proposed in [3] [4].

1. Objectives

This paper aims to describe and discuss the security model implemented as part of TRENCADIS. The security services of grid are not altogether different from those of other distributed system paradigms. Specifically, an effective security model must ensure a set of security primitives: Identity verification (authentication), authorization, access control, data integrity, data confidentiality and availability.

TRENCADIS architecture combines the security requirements in three main objectives:

- Secure authentication and communication.
- Management of VO security policies.
- Confidentiality and privacy control of medical data.

[1] TRENCADIS: Towards a GRid ENvironment to proCess and shAre DIcom objectS. TRENCADIS means “mosaic” on the language of Valencia Region.

In the next points the TRENCADIS security model is described and discussed. First, a general view of the TRENCADIS architecture is presented in which the security model is applied to accomplish the three security aspects proposed. After that, the security model and the components required in the deployment of the architecture are covered.

2. Related Work

2.1. Secure authentication and communication

Almost all Grid components and Grid Middlewares use the Grid Security Infrastructure (GSI) for authentication [7]. GSI also provides mechanisms for secure communication. These mechanisms deal with the security related aspects of connection establishment as well as message protection. Currently, the message protection could be enforced at two different levels: integrity, guaranteed by the message signature, and privacy, guaranteed by the message encryption and signature.

2.2. Management of VO Security Policies

This objective refers to the need of authorization and access control policy mechanisms at VO level.

The TRENCADIS middleware must include a fine-grain security access control for Grid services and resources. The system must consider the agreements reached between the VO and the different organizations which comprise the VO for the management of the shared computing resources.

Virtual Organization Membership Service (VOMS) has long been proposed as a solution to this problem [8]. VOMS utilizes an extended X.509 certificate specification for defining extra attributes. VOMS defines groups, roles and capabilities. Combinations of the names of these serve as attributes for users [9].

2.3. Confidentiality and Privacy Control of Medical Data

As medical data may be stored for years, long-term storage of encrypted data seems to be a reasonable approach for protecting confidential data from being exposed. Different organizations (virtual or not) need to combine their efforts in a collective purpose of controlling data privacy.

The combination of the data controllers in a scheme for managing decryption keys protects data from disclosure [10].

Different administrative domains agreed to contribute with the collective purpose of protect data from being exposed by users granted with physical or administrator access. The complexity of disclosing protected objects requires compromising the security of a certain number of services deployed by completely different administrative domains.

A previous work have introduced the idea of using VOs as a natural way of define a key sharing schema in a Grid environment, considering each VO as a different administrative domain with the responsibility of guard key shares and object owners being enabled to decide the trusted VOs they will use for key sharing [11].

3. TRENCADIS Architecture

The architecture defined in TRENCADIS project is horizontal and can deal with in different type of objects of the DICOM standard (medical images, structured reports, waveforms, etc…).

In particular, this paper focuses on the Security model applied above the architecture. This section briefly explains the different layers defined in the TRENCADIS architecture and the features associated to it. TRENCADIS is a Service-Oriented Architecture (SOA) in which the usage of resources is represented as Grid services.

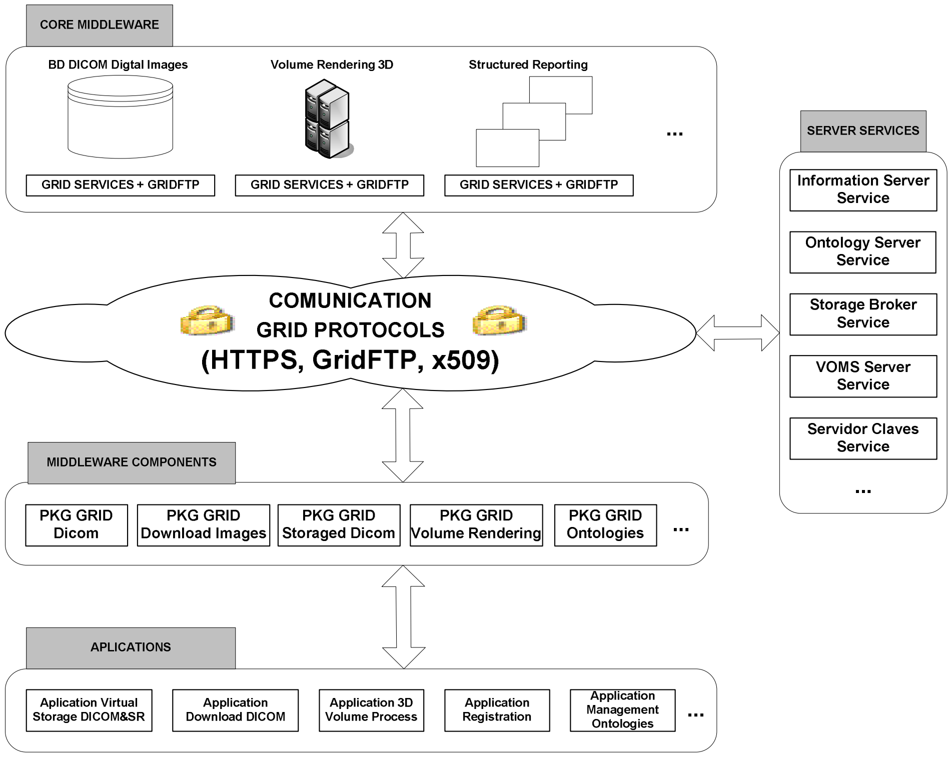

Figure 1. General Scheme of TRENCADIS Architecture.

As described in Figure 1, TRENCADIS comprises five layers:

- *Core Middleware Layer*. This layer provides with the basic resources (Databases, High-Performance Computers, etc…) of the environment, which are offered as services using well defined and standard interfaces in Web Services Definition Language (WSDL), using protocols such as Simple Object Access Protocol (SOAP) or GridFTP and data formats like XML Schemes. It provides upper layers with a unique interface to all resources of the same type. For example, a DICOM storage of structured reports can be implemented using relational databases or a plain directory in a hard disk, but the interaction interface will be the same. The Services of this layer are part of the infrastructure.

- *Server Services Layer*. This layer defines the services that implement server tasks, such as the location of resources. This layer interacts directly with the services from Core Middleware and Server Services Layer and with the upper-level components. The Services of this layer are part of the infrastructure.
- *Communication Layer*. It defines the protocols that will be used by the services that have been implemented in the lowest layers (Core Middleware and Server Services). GridFTP protocol is used for transferring a large amount of data, and SOAP above HTTPS is used for interacting with the services.
- *Components Middleware Layer*. This layer contains the highest components that interact with services of the Core Middleware and the Server Services Layer. These components provide the applications with an object-oriented interface for the development of the applications for managing, processing and sharing DICOM objects in general.
- *Applications Layer*. This layer comprises the applications for managing, processing and sharing DICOM objects.

4. Infrastructure of Security in the TRENCADIS Architecture

This section describes how TRENCADIS is architected to realize the security goals discussed in the previous sections. The different security services and components of the infrastructure are briefly explained, as well as the exposed interfaces.

All the services presented in this Section were implemented and deployed as part of the *Infrastructure Layer*, a virtual layer that groups the *Core Middleware* and the *Server Services* layers.

4.1. GSI Services

TRENCADIS uses the GSI for enabling secure authentication and communication over an open network. Since the core of GSI is the use of X.509 certificates, an international standard for public key infrastructure (PKI), a set of Grid services is required to manage the issuing and revocation of PKI credentials for services and users.

GSI minimally requires a basic PKI certificate management infrastructure (request, registration and revocation). TRENCADIS deploys the Certificates Management Service in the Server Service Layer, that serve as entry point for new PKI certificates requests managed by one or more Certificate Authority (CA). It is possible to use any of the existent CA implementations (OpenCA, Microsoft Certificate Services, etc).

The Certificates Management Service handles new certificate requests for users and resources which want to be integrated in a specific Grid infrastructure. Registration and revocation decisions must be managed by the authorized operator of the administrative domain of the CA.

The service exposes two different interfaces: a client interface and an administration interface available only for the service administrator.

The client interface exposes the following operations:

- **SendRequest:** It sends the users and resources new certificates requests to the CA.
- **RetrieveCertificate:** It retrieves the certificates issued by the CA.

The administration interface exposes the following operations:

- **SignRequest:** It signs the new certificates requests.
- **RevokeCertificate:** It revokes certificates.

The Communication Layer relies on the Globus Toolkit to provide the client and the server SOAP/HTTPS and GridFTP protocols.

4.2. VOMS Services

Different real domains could be managed as a single administrative domain in TRENCADIS. The architecture supports VO management throughout VOMS.

The VOMS Service is implemented in the Core Middleware and in the Server Service Layers. The components of the service are very similar in both layers. The main difference between the components implemented in each layer is that the Core Middleware Layer requires an additional component to facilitate the interaction with devices.

The VOMS Service manages user memberships, roles and capabilities in one or more VOs. The service is exposed in two different interfaces: a client interface and an administration interface available only for the service administrator. Both interfaces enable a Web client application for the service.

The administration interface exposes the following operations:

- **CreateVO/DeleteVO:** It creates/deletes a VO.
- **CreateUser/DeleteUser:** It creates/deletes a user in the VO.
- **CreateGroup/DeleteGroup:** It creates/deletes a group in the VO.
- **AssignGroup/ResignGroup:** It registers/deregisters a user in a group.
- **CreateRole/DeleteRole:** It creates/deletes a role in the VO.
- **AssignRole/DismissRole:** It assigns/dismisses a role to/from a group.

The client interface exposes the operation RequestProxy that returns a VOMS proxy certificate for a user identified by a user certificate.

The Grid services deployed in the layers Core Middleware and Server Service require a component which reads the attributes list from the user credentials and evaluates a local access control policy returning a decision. The decision could be to permit or to deny the execution of a request submitted to the service. This component is named the Gatekeeper.

In a typical TRENCADIS usage scenario, a subject (e.g. user, service) wants to take some action on a particular virtual storage DICOM. The subject submits its query to the Storage DICOM Service protecting the resource, which examines the request, retrieves the ontologies that are applicable to this request and the subject's VOMS attributes, and contacts a Gatekeeper component that determines whether access should be granted according to the policy rules for evaluating VO groups' access to ontologies.

In this context, technologies for access control and enforcement policies (e.g. SAML, XACML) could be used for exchanging authentication and authorization data between security domains. However, this version of TRENCADIS lacks this functionality for simplicity. Future releases of the architecture were expected to be compliant with some standard access control policy language.

Besides to evaluate decision chains, the Gatekeeper implements a set of operations in the Core Middleware Layer. The Gatekeeper automatically registers authorized users in the grid-mapfile when needed (for example for the GridFTP protocol) and delete it when the requested operation ends.

The Gatekeeper components in the Server Service Layer could require further operations with the purpose of allowing alternative authorization methods.

4.3. Privacy Services

Often, production Grids comprises a few VOs. In real applications, Grids deployed by a single VO are more commonly found that one can expect. In such scenarios, it is not possible to share keys over VOs with a reasonable security level because there are no enough key stores available in different administrative domains.

The VO-based definition of administrative domain limits the generalization of the model. Therefore, considering the experiences acquired by the HealthGrid community throughout the development of a middleware Grid for managing DICOM objects, the concept of the administrative domain is updated in this work, preserving the rest of the architecture.

Parties involved in the deployment of Grid services for storing DICOM objects fits two different categories: organizations feeding the Grid with their own data, and organizations using the data provided by others. Data providers form the group of organizations interested in guaranteeing the confidentiality and privacy of the data. Consequently, administrative domains should be defined as individual organizations (virtual or not) which contribute with their own data to the Grid and which are identified as data controllers with the responsibility of protecting data from unauthorized use.

In practice, different administrative domains can be identified using PKI. Each organization (virtual or not) participating in the sharing scheme must contribute with a CA to the Grid, and must deploy their own key sharing Grid services exposing credentials signed by their own CA. The objects owners' will be enabled to discover administrative domains, properly identified, and to distribute the decryption keys over trusted participants.

The information required to identify the administrative domains that could be combined to retrieve the decryption key is stored in the header of the encrypted object. However, this information is not enough to rebuild the key. Besides, each administrative domain needs to keep track of the object identifiers and its corresponding key share.

The object identifier should provide the ubiquity property to the encrypted object. That is the capability of been physically present in more than one operation or storage element without changing its identity.

A 128-bits Encrypted Object Unique Identifier (EOUID) is randomly generated by a Grid service and assigned to each new encrypted object stored in the grid. The EOUID could be retrieved with a cost-less operation from the header of the encrypted object. All key sharing services references use this identifier.

Privacy Services require the definition of two new components in the infrastructure. The Key Sharing Service and the EOUID Generator Service are implemented in the Server Service Layer.

The Key Sharing Service keeps the decryption key shares and further information related to the encrypted objects as the EOUID and the Message Authentication Code

(MAC) signature for the encrypted object. This service provides to the authorized clients the data needed to decrypt objects in a secure way.

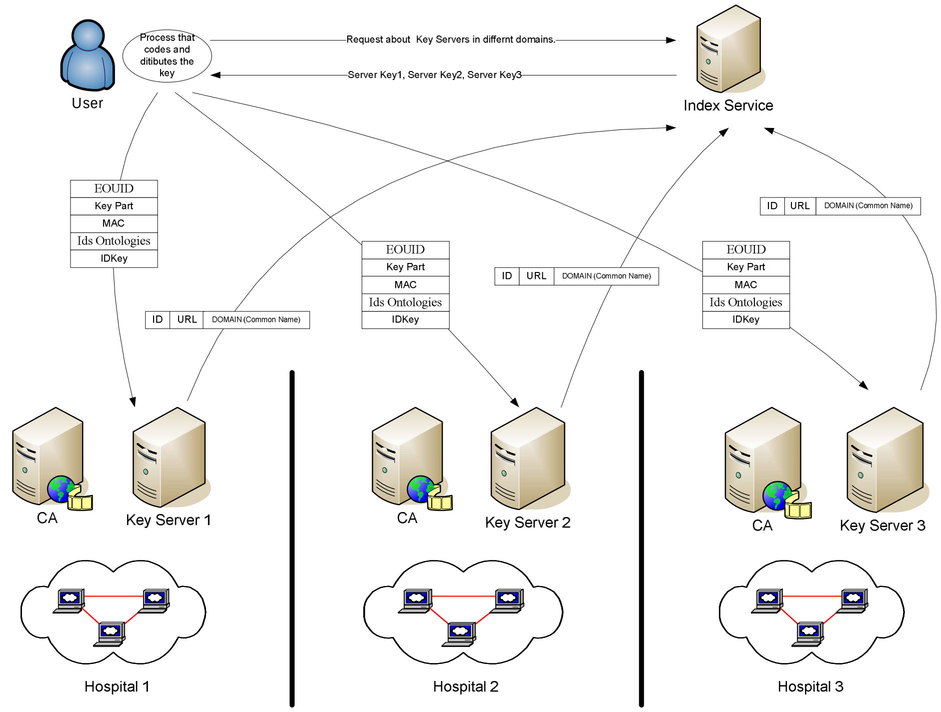

Figure 2. Publication of Key Servers in different administrative domains and distribution of the keys.

Clients need a valid VOMS proxy certificate for executing operations in the service. The service presents the attributes contained in the proxy to the Gatekeeper which enforces the access control policies of the domain. The Gatekeeper could use supplementary information managed by external Grid services to authorize a request.

A work describing Grid middleware that uses an ontological schema to create virtual communities and to define common targets for sharing, transferring and processing DICOM medical images in a distributed environment [6] and a work describing Grid middleware for managing DICOM Structured Reporting Objects [3], have been recently adopted by the medical community. Both projects are based in the TRENCADIS software architecture, and both implement a Gatekeeper component that makes use of the ontology information used for creating virtual storages of DICOM objects to control the access to the virtual storages. Ontologies are managed by an Ontology Server which provides the links between ontologies and VOMS attributes that serve as access control policy.

The Key Sharing Service operations are exposed in a single interface:

- **xmlGetError:** It returns the last error which occurred in the service.
- **xmlSaveSubkey:** It stores a decryption key share in the key server.
- **xmlRetrieveSubkey:** It retrieves a decryption key share from the key server.
- **xmlUpdateKey:** It updates an existent decryption key share.

The EOUID Generator Service produces valid EOUID for a new encrypted object. Although the operations of this service are very simple, it is a good idea to deploy several instances of the service in each administrative domain, because this service could be easily targeted by a malicious client who wants to produce deny of service failures. We propose to deploy one EOUID Generator Service per each Key Sharing Service deployed in the administrative domain.The EOUID Generator Service operations are exposed in a single interface:

- **xmlGetError:** It returns the last error which occurred in the service.
- **xmlGetEOUID:** It returns a new valid EOUID.

5. Conclusions and Future Plans

In this work, a security model for the TRENCADIS architecture has been described. A suitable security level is obtained by defining a set of services and components in the infrastructure (Core Middleware Layer and Server Services Layer). The model aims three main objectives: a) Secure authentication and communication; b) Management of VO security policies; and c) Confidentiality and privacy control of medical data.

The security model is flexible and expansible, and could be integrated with any of the functionalities provided with the architecture through components in the Components Middleware Layer. All general security services have been specified and deployed with the infrastructure, in the same way the additional Gatekeeper component required by new services has been described.

Nowadays, TRENCADIS offers a high-level object-oriented interface with a given functionality [3] [4] (create Virtual Storages, Download DICOM objects, Upload DICOM objects, etc…) through the Components Middleware Layer, and it increases the productivity of code developers for building applications for managing DICOM objects in a secure way. Our focus at this moment is to integrate the security model with the actual functionalities.

A practical deployment of TRENCADIS is the CVIMO architecture. CVIMO (Valencian Cyberinfrastructure for Medical Imaging in Oncology) is a project funded by the regional government of Valencia (Conselleria d'Empresa, Universitat i Ciencia de la Generalitat Valenciana, code GVEMP06/004) involving 5 hospitals of the Land of Valencia interested on sharing Medical Images on the fields of Liver, Lung and Central Nervous System Cancer. This three virtual communities have been deploy and templates for Structured Reports (both for staging and follow-on), following the DICOM standards have been developed. A tool for filling-in the Structured Reports has been developed.

A use case is as follows. When an interesting case is detected, the radiologist sends it from the workstation to the CVIMO local storage of the Hospital through DICOM protocols. Automatically, the system registers this case in the Global Virtual Storage and makes it available as "non-informed" case. Then, the radiologists can fill-in the structured report through a web-browser, and save this report on the local system. Then, both image and structured report are made available for sharing within the VO and indexed through the searching criteria of the ontology in which the structured report was made, considering the information filled-in on the report. Any radiologist with a valid VOMS certificate belonging to an authorized VO and group can request the downloading of the study to the local machine.

6. Acknowledgements

The authors wish to thanks the financial support received from The Spanish Ministry of Science and Technology to develop the project “Componentes de Nueva Generación para la Explotación Eficiente de Infraestructuras Grid en E-Ciencia”, with reference TIN2006-12890.

References

[1] Digital Imaging and Communications in Medicine (DICOM) Part 10: Media Storage and File Format for Media Interchange. National Electrical Manufacturers Association, 1300 N. 17th Street, Rosslyn, Virginia 22209 USA.

[2] “DICOM Structured Reporting”. Dr. David A. Clunie. ISBN 0-9701369-0-0. 394 pages softcover.

[3] I. Blanquer, V. Hernandez, and D. Segrelles. “TRENCADIS – a WSRF Grid MiddleWare for Managing DICOM Structured Reporting Objects”. Proceeding of HealthGrid 2006. ISBN. Technology and Informatics. Ed. IOPress. ISBN:1-58603-617-3

[4] I. Blanquer, V. Hernández, J. D. Segrelles. “TRENCADIS – A Grid Architecture for Creating Virtual Repositories of DICOM Objects in an OGSA-based Ontological Framework”. Lecture Notes in Computer Science (LNCS), Subseries Lecture Notes in Bioinformatics (LNBI). Biological and Medical Data Analysis. ISSN: 0302-9743, ISBN-10: 3-540-68063-2, ISBN-13: 978-3-540-68063-5

[5] I. E. Magnin, J. Montagnat. “The Grid and the Biomedical Community: achievements and open issues” EGEE User Forum, CERN, Geneva, Switzerland, March 1-3, 2006

[6] I. Blanquer, V. Hernández, J. D. Segrelles. “An OGSA Middleware for managing medical images using ontologies”. Journal of Clinical Monitoring and Computing. 2005 Oct;19(4-5):295-305

[7] Grid Interoperations Cook Book. https://twiki.cern.ch/twiki/bin/view/EGEE/InteropsCookBook

[8] R. Alfieri, R. Cecchini, V. Ciaschini, L. dell’Agnello, Á . Frohner, A. Gianoli, K. Lörentey, F. Spataro. “VOMS, an authorization system for virtual organizations”. In Proceedings of the 1st European Across Grids Conference, 2003

[9] A. Frohner, V. Ciaschini. “VOMS Credential Format”, EDG Draft, 2004

[10] L. Seitz, J. M. Pierson, L. Brunie. “Encrypted storage of medical data on a grid”. Methods of Information in Medicine. 2005; 44(2):198-201

[11] E. Torres, C. De Alfonso, I. Blanquer, V. Hernandez. “Privacy Protection in HealthGrid: Distributing Encryption Management over the VO”. Proceedings of HealthGrid 2006. ISBN. Technology and Informatics. Ed. IOPress. ISBN:1-58603-617-3