Handling confidentiality and privacy on Cloud-based Health Information Systems

Carlos Serrão, Elsa Cardoso

[1]ISCTE- Instituto Universitário de Lisboa, Av. das Forças Armadas, 1649-026 Lisboa, Portugal

{carlos.serrao, elsa.cardoso}@iscte.pt

Handling confidentiality and privacy on Cloud-based Health Information Systems

ABSTRACT

Health-related data includes not only personal information about the patient, but also specific information about the patient health problems, supplementary diagnostic exams results, and much more. All of this information is extremely sensitive and contain private personal information that should only be accessed by the proper entities and actors, for special specific purposes.

Throughout this article the authors propose and describe an approach to address security and privacy of health-related data based on rights management technologies and present an architecture that will be able to minimize the security risks and the privacy concerns that were previously identified. The approach followed in this work consists in the reutilisation of an open-source and open-specifications rights management system, designing and adapting the necessary components to address the specific security and privacy requirements that need to be faced when managing health and patient data.

KEYWORDS

EHR, PHR, rights management systems, security, confidentiality, privacy

INTRODUCTION

The evolution of Information and Communications Technology (ICT) has created opportunities for the development of new products and services, optimisation of organisational processes, and collection, integration and better interaction with large amounts of data from multiple sources. This evolution has impacted on many different aspects of people lives, marking the evolution of our society and making us more and more dependent of this IT-based infrastructure (Haluza & Jungwirth, 2015).

Both entities and individuals are surrounded by technology that silently captures all kinds of data about them taking advantage of smaller-sized devices (Zheng et al., 2014). Data is travelling across multiple domains, from well-known and centralised data centres to decentralised and virtualised warehouses, commonly referred as "clouds", where it is stored and processed in ways never imagined before. Different types of organisations are taking advantage of these new possibilities, as a way to optimise processes efficiency, integrate information, provide better services, reduce costs and much more (Thilakanathan, Chen, Nepal, Calvo, & Alem, 2014). The health sector is also embracing these changes and increasingly adopting advanced decentralised ICT systems, to improve service efficiency and quality - this new breed of health information systems (HIS) are also identified as Electronic Health Record (EHR) systems as well as the Personal Health Records (PHR) (Brennan, Downs, & Casper, 2010).

The PHR is a kind of health record where the health patient and care data is maintained by the patient itself. This PHR can include lab results (like supplementary diagnostic means), data provided by smart health and activity monitoring devices, or even by smartphones. The purpose of this PHR is to provide a more accurate medical history, when combined with other sources of information, such as the EHR that contains data handled by institutions and entered by clinicians. Both PHR and EHR are represented in digital format, containing important health data about one or more patients (Cushman, Froomkin, Cava, Abril, & Goodman, 2010).

A trend that is also being followed by the health sector is the continuing systems integration and the migration of data and services to the cloud (Haux, 2006). This is an important change because it affects the direct relation with the information, the way and places where it is stored, and the potential threats that it might be subject to (Juliadotter & Choo, 2015). The introduction of these systems, although extremely important to the improvement of a patient-centred care, originated huge concerns about information privacy and confidentiality, in particular due to the sensitivity of the information at stake (Rodrigues, de la Torre, Fernández, & López-Coronado, 2013).

Cybercriminals are increasingly directing their attacks towards health and medical data, offering them additional opportunities to commit fraud or embrace other similar criminal schemes. Data breaches

occurring in the healthcare industry can have a financial and reputational effect, but can also have dramatic effects for the patients due to the nature of the disclosed data (Fernando & Dawson, 2009). Patient's identity and any other associated medical information could be stolen directly from hospitals, healthcare insurance companies, and from any system that manages medical records in a digital format.

The purpose of this work is to present an approach that is based on rights management systems as a way to improve the confidentiality and privacy of PHR and EHR, helping the development of a secure and trustworthy environment that enables entities to access and share digital clinical information in a governed way. This article starts by providing an introduction to the problem, followed by some information about the major confidentiality and privacy requirements to address in the health-related data field, and also some related work on the existing literature. Following this introductory stage, the proposed approach is presented, describing the architecture and some of the security mechanisms that are used to fulfil the confidentiality and privacy needs. Finally, some conclusions, final remarks and further research directions are presented.

HEALTH INFORMATION CONFIDENTIALITY AND PRIVACY REQUIREMENTS

Every single day, our digital existence faces menaces from different threats that are escalating in terms of sophistication. Data breaches and leakages have the potential to expose several millions of records that can be used by criminals to conduct all types of illegal activities (Choo, 2011). Medical and health data is considered extremely sensitive information, deserving the cyber criminal's special attention (Meyer & Pyles, 2005).

A recent study revealed it is possible to conclude that this is already a serious problem with a growing trend. The study (Redspin, 2014) reported that nearly 30 million Americans have had their personal health information breached or accidentally disclosed since 2009, and also that in 2013, the number of major data breach cases of medical records, also called protected health information (PHI), was 804,

affecting over 29.2 million patient records (Redspin, 2014). These breaches have not only a financial impact but also a reputational one.

Several risks and threats can be identified and categorised when considering health-related information (Juliadotter & Choo, 2015). These threats can have an external, internal, intentional or non-intentional origin, such as human, natural or environmental, or technological threats (Djambazova, Almgren, Dimitrov, & Jonsson, 2011). All of these threats need to be handled when considering the medical and patient information and the ways to protect this vital information.

Currently, there are several legislative initiatives created to deal with the confidentiality and privacy requirements of the health and medical information. The US has established rules for accessing, authenticating, storing, auditing and transmitting medical information, called HIPAA (Health Insurance Portability and Accountability Act). HIPAA (Wu, Ahn, & Hu, 2012) protects the patient information, present in electronic medical records (named as Protected Health Information - PHI) namely the information that doctors and nurses input into the electronic medical record, recorded conversations between doctor and patient and also financial information (Wu et al., 2012). HIPAA also defines how much information can be disclosed and who can access such information, as well as the patient rights over the information and the way this information can be shared. The European Union has also established some more generic initiatives from the Council and European Parliament, such as the "The Data Protection Directive" (Birnhack, 2008), that protects the processing and free movement of personal data, including the health-related data.

All of these legislative initiatives are quite relevant. However, the interoperability and exchangeability of health-related information are also quite important. Especially, in a digital world where the benefits from a decentralised and interoperable electronic health information system, can be used to create better processes and provide better health services - a vision of an integrated health information system, completely decentralised, located on the cloud (Nepal, Ranjan, & Choo, 2015). The adequacy of these legislative initiatives with the requirements of this system needs to be aligned (Massey, Otto, & Antón, 2008).

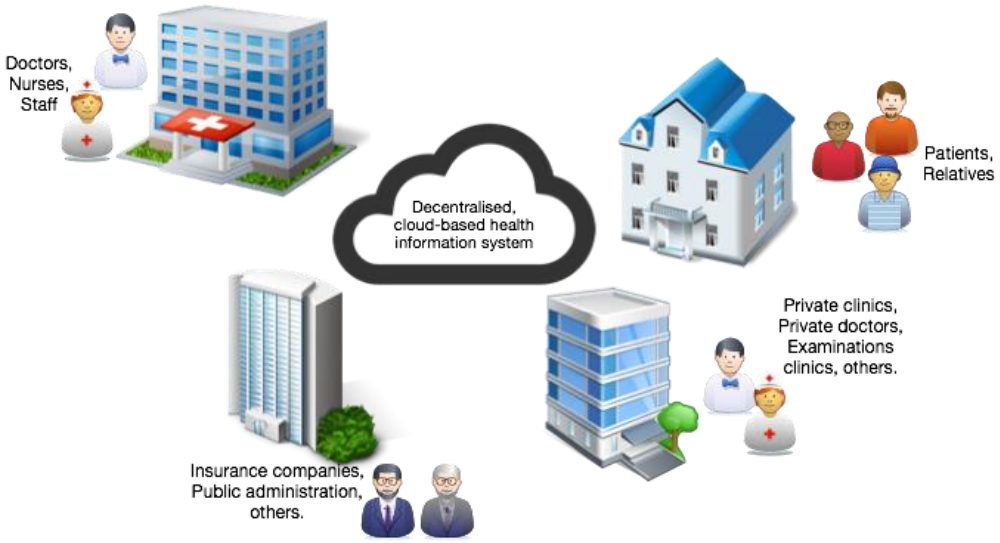

Figure 1. Decentralised cloud-based health information system

This ideal architecture, as depicted in Figure 1, is the one in which a set of institutions (hospitals, private clinics, insurance companies, medical examination institutes, public administration, public health-related institutions, and others) and actors (doctors, nurses, administrative staff, patients, relatives and others) can share a common vision of the patient medical data (including health-related information as well as financial and administrative information). The vision of a cloud-based health information system could facilitate the interaction between the different institutions and actors (including the patient), in a way that would provide benefits to all of them.

The problem with such architecture is the fact that the information on this cloud-based information system, should be kept in a secure and private way. Only authorised entities would be able to access certain parts of the information, on a certain time, in order to conduct specific operations. This system should encompass the possibility to define and uphold an information governance architecture, in which it would be possible to specify the specific access rights for institutions and actors. The proposal for this system in this paper is based on a rights management system.

RELATED WORK

The issue of confidentiality and privacy on electric health data is not new. Many researchers and companies have been working on mechanisms and systems to offer these two important requirements

on health information systems, especially in cloud environments. Some of the authors (AbuKhousa, Mohamed, & Al-Jaroodi, 2012; Juliadotter & Choo, 2015; Rodrigues et al., 2013; Sultan, 2014) are focused on the specific security and privacy requirements of cloud-based electronic health record systems, where they analyse the aspects of the security and privacy maintenance on HER, such as authorised access to data, data confidentiality, patient's consent to allow others to access to data, data relevance, information ownership, information consistency, audits (Wang, Li, & Li, 2012) and archiving. Also some more ethical, legal and standardization frameworks are necessary to implement interoperable environments on e-health systems, to enable the exchange of information between the different stakeholders, are also referred by some other authors (Benson, 2012; Birnhack, 2008; Brailer, 2005; Haux, 2006; Massey et al., 2008; Saaranen, Parak, Honko, Aaltonen, & Korhonen, 2014; Wu et al., 2012).

On the specific aspect of the mechanisms existing to offer the necessary protection to the confidentiality and privacy of the e-health information on cloud environments, there are two state-of-the-art common approaches: cryptographic and non-cryptographic approaches (Abbas & Khan, 2014). In the cryptographic approaches, the authors refer to the usage of Public-Key Encryption hybrid approaches, Secret-Key Encryption approaches and Alternative Cryptographic primitives, such as Attribute-Based Encryption (M. Li, Yu, Zheng, Ren, & Lou, 2013) or homomorphic encryption (Ikuomola & Arowolo, 2014). Non-cryptographic approaches mainly use certain policy-based authorization infrastructure that allows the data objects to have access control policies.

Also, new approaches are highlighted in the literature, upholding the patient control of its own health data on the cloud as a way to control the access from different stakeholders, authorizing them through the setting of an access tree supporting flexible threshold predicates (Zhou, Lin, Dong, & Cao, 2015).

More recent approaches also include the usage of block chain techniques as a way to improve the privacy of security on the cloud (Lazarovich, 2015). Block chain addresses the concerns of security, scalability and privacy of electronic medical records.

Some other authors (Jafari, Safavi-Naini, & Sheppard, 2011) also refer the usage of digital rights management to establish a patient-centric DRM approach to protect privacy of health records stored in

a cloud storage based on the patient's preferences and without the need to trust the service provider. The usage of rights management systems is an approach that is not that explored in the literature as a way to provide confidentiality and privacy to health information. The approach in this article differs in the sense that it presents not only the application of an open and standards based rights management framework but also describes also the necessary processes to preserve the confidentiality and privacy of health data.

RIGHTS MANAGEMENT SYSTEMS TO PROVIDE CONFIDENTIALITY AND PRIVACY OF HEALTH INFORMATION SYSTEMS

Rights management systems, also referred as digital rights management (DRM) systems started being used in a broader way mainly by the music and book industry as an attempt to prevent piracy. However, public (and in some cases, artists and authors) acceptance was so negative, that users started regarding it as a "bad thing". The music industry, and also the movie, book and game industry, looked at DRM as a way to prevent non-authorised replication of copyrighted material – mainly as a copy protection mechanism (for instance as a way to prevent CDs or DVDs from being copied). However, rights management systems should be regarded as something that is much more than simply providing copy-protection mechanisms. A rights management system can also be seen as an information governance tool that can be used to control fine grained access to information, for instance to safeguard the confidentiality and integrity of that information (Sujansky, Faus, Stone, & Brennan, 2010).

Some of the existing rights management solutions in the market are both closed-source and commercial. These solutions tend also to be vertical in the sense that they are developed having into an account a specific business model, a specific type of content and a particular type of device (this is for the instance the case with Windows Media DRM and Apple iTunes DRM systems). Although these solutions represented an apparent solution for the media industry, these systems were not interoperable, meaning that content protected and governed by a system could not be used by other. To address the interoperability problems, more open rights management solutions have emerged, most of them resulting from the combined efforts of academia or industry consortia: OMA-DRM (Open Mobile

Alliance DRM) (Buhse & Van der Meer, n.d.), OpenIPMP (Open Intellectual Property Management and Protection) (OpenIPMP, 2003), DReaM (DRM everywhere available) (Gerard Fernando, Tom Jacobs, & Vishy Swaminathan, 2005), MIPAMS (Multimedia Information Protection and Management System) (Victor Torres, Delgado, & Llorente, 2006), Marlin (Marlin, 2006) and OpenSDRM (Open and Secure Digital Rights Management) are some of the examples of rights management solutions that try to address interoperability. These systems use standard-based architectures to offer an interoperable content governance architecture that enable the content protection and governance and rights management interoperable. Interoperability is a key issue for content protection and governance, and in consequence, for the control of privacy and confidentiality of health-related information (Brailer, 2005; Liyanage, Krause, & de Lusignan, 2015; Saaranen et al., 2014) to allow the exchange of patient information between different highly distributed, cloud-based, health information systems (Benson, 2012).

Having into consideration the interoperability requirements of health information on information systems, choosing an interoperable rights management platform was one of the major selection criteria. Additionally, other characteristics were also fundamental, such as the security, the standards compliance and its openness, both in terms of software availability, documentation and source code access. Among the different rights management solutions identified before, all of them satisfied the interoperability criteria and were based on open standards. Rights management solutions such as OpenIPMP and DReaM were not considered for evaluation and usage because they are no longer active as rights management projects. OMA-DRM, MIPAMS, Marlin and OpenSDRM rights management platforms are solutions that were developed with interoperability as a major concern and were based on open standards. Both OMA-DRM and Marlin possess open specifications, but although there exist specific implementations, none is available in open-source format and therefore is hard to implement a rights management solution to address confidentiality and privacy without the needed access to software source-code. MIPAMS and OpenSDRM are two rights management solutions that emerged in the academic background, also based on the interoperability and open standards principles. Both of them also use an service-oriented architecture to enable the decoupling of the different services offered by

each of the rights management frameworks, enabling the adaptability to different usage scenarios and possible business models. MIPAMS and OpenSDRM describe and document the offered services interfaces and, on top of that, OpenSDRM is completely open-source while MIPAMS offer also some parts of the solution in open-source regime.

Considering the multiple existing rights management solutions existing and the specific interoperability requirements in terms of health-related data confidentiality and privacy, the approach followed by the work described in this article was based on the usage of the OpenSDRM open rights management system (Carlos Serrão, 2008). OpenSDRM is an open and distributed rights management architecture that allows the implementation of different information governance models. OpenSDRM was developed considering interoperability aspects (Carlos Serrão, Rodriguez, & Delgado, 2011) with a modular design that allows the composition and reconfiguration of the system to allow interoperability (Carlos Serrão, Dias, & Kudumakis, 2005; Carlos Serrão et al., 2011) with other non-rights management components, through well-defined interfaces using a service-oriented approach. The system was also developed considering the scalability and adaptability to different information governance scenarios (Víctor Torres, Serrão, Dias, & Delgado, 2008). Another important aspect that makes OpenSDRM a good candidate solution for the governance of health information confidentiality and privacy is the fact that it has already been used to implement different content governance cases such as digital music e-commerce services (C Serrão, 2005), controlled access to video-surveillance data (C Serrão, Serra, Fonseca, & Dias M, 2003), e-commerce and controlled access to large earth observation products (Carlos Serrão & Dias, 2002) and home networks music jukeboxes (Carlos Serrão, Serra, Dias, & Delgado, 2006).

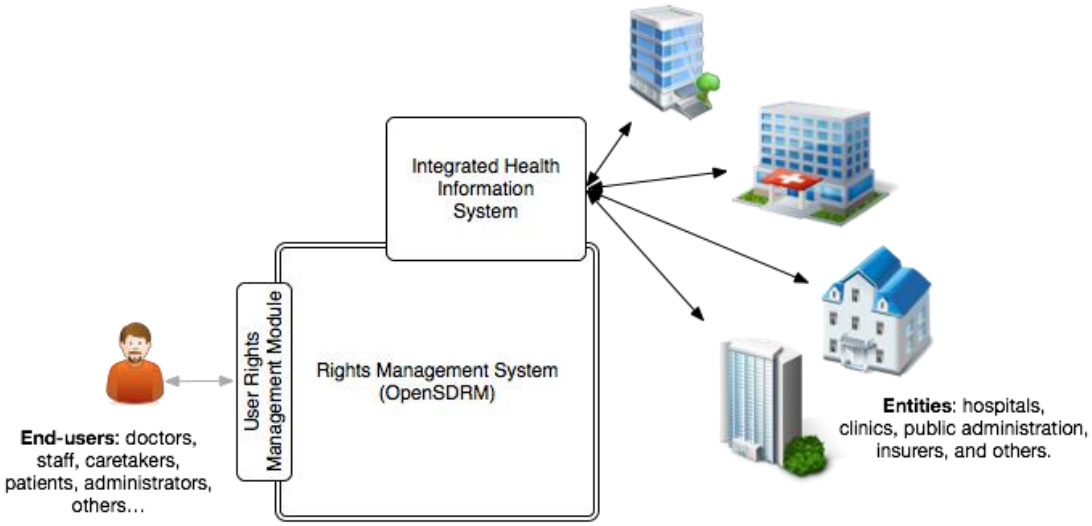

Figure 2. Overview of the rights management system integrated with the health information system

For the proposed scenario (Figure 2), the cloud-based health information system (or at least the different functions that are provided by it) will need to be integrated with the rights management system - OpenSDRM - using the different methods. The Rights Management System (RMS) offers an open web-based API (Application Programming Interface) that enables the integration of any external system with it. In this case, the cloud-based Health Information System (HIS) can use that web API to invoke the necessary information governance operations that are necessary.

In this system it is also important to notice that the different actors will have to use a specific module, the User Rights Management Module (URMM) that is responsible for the enforcement of the information governance rules at the end-user side - for the different scenario actors.

The rights management system, OpenSDRM, is an open rights management framework that is composed by a set of different services (Figure 3). Due to its distributed and decoupled nature, the RMS implements services on the server-side as well as services on the user-side. On the user-side, the authorisation service is responsible for handling the requests to access some type of information or content on the user device (whatever that device might be, as long as it is integrated with the RMS), processing the requests and matching them to existing access credentials, licenses and permissions to perform a given operation over the information or associated content. Also, on the end-user side the information rendering service is responsible for verifying the necessary requirements to perform a

requested operation over the governed information or content (such as encryption, scrambling, or others) and effectively allowing or denying the end-user requested operation on that information or content.

The larger part of the rights management responsibility is entered on the server-side - a set of components with a well-defined API that allows the integration between the necessary ones to implement the specific information governance model. These services are:

- Information storage and distribution service: this service is responsible for the storage and distribution of governed information and content in a protected manner;
- Information protection service: the service is responsible for the information and content protection. Any information and associated content is protected by specific protection tools and mechanisms that may change according to the information, content and the implemented governance model;
- Information registration service: this service is responsible for registering the information and associated content on the platform that will be used to uniquely identify it on the system. This unique identifier is used to identify the governed information and content throughout the entire system;
- Payment service: if the governance model includes the possibility to build some type of information or content trading, this payment service is responsible to communicate with a payment gateway that implements the necessary mechanisms to process payments - this service, although part of the RMS will not be used in the HIS scenario;
- Protection tools service: this service is responsible for the registration of protection tools on the system and for making those tools available to the information protection service to use when implementing the information protection schemas (such as encryption, scrambling, watermarking and others);
- Authentication service: handles the registration of users and services on the system as well as the requests for authenticate users on behalf of other services;

- Licensing service: this is one of the most important services of the rights management framework, responsible for creating license templates, define and produce new content licenses (that represent the type of rights, permissions and/or restrictions of a given user, or group of users, over the governed information and associated content) and provide licenses, upon request, to specific users.

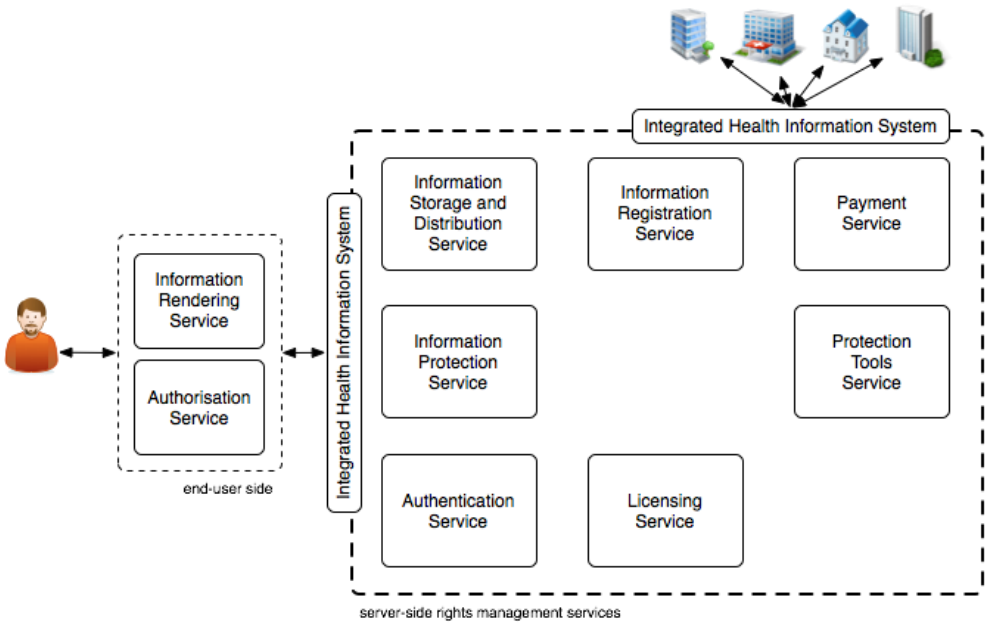

Figure 3. Detail of the rights management services integrated with the HIS

These RMS services will be used to implement the necessary mechanisms to govern the health-related information and any associated content on the HIS. The next sections of this paper will describe some of the mechanisms used to ensure the necessary confidentiality and privacy requirements.

SERVICES AND USERS/ACTORS REGISTRATION ON THE PLATFORM

The approach to this platform requires a trustworthy system that requires all the necessary systems to be registered on that platform. Each one of the different services on the RMS, have to be registered in the platform. This registration process assigns unique credentials to each one of the services, ensuring that they are uniquely registered and that these credentials will be used to identify and differentiate the services in future interactions (Figure 4).

This registration process is conducted by the authentication service (AS) that issues credentials to all the other services and acts as a central trust mechanism, delegating its trust to all the different services on the system. All the communication channels between the services is handled over a secure and authenticated channel, using Secure Sockets Layer/Transport Layer Security (SSL/TLS) – ensuring the authentication and security of the hosts where the services are deployed and allowing the establishment of secure communication channels (Thomas, 2000).

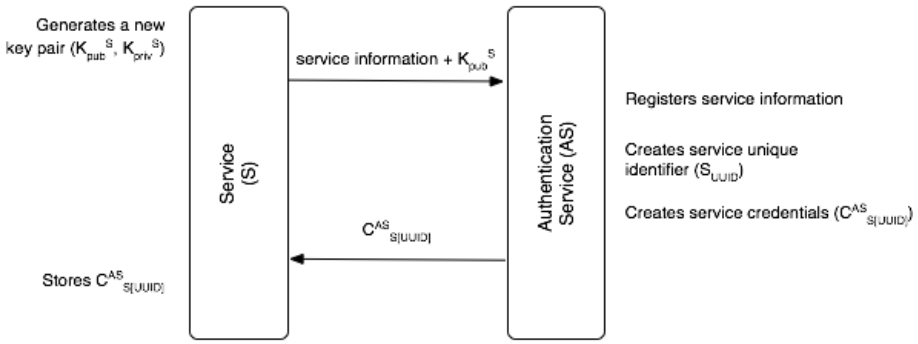

Figure 4. Registration of the services on the platform

1. The authentication service (AS) has already a key pair ($K_{pub}^{AS}, K_{priv}^{AS}$) and a self-issued ($C_{AS}^{AS}$) credential or issued by other trustworthy entity (C_{AS}^{CA}) - for instance some certification authority (either private or public);

2. The service (S) that needs to be registered generates a key pair ($K_{pub}^{S}, K_{priv}^{S}$) and sends a registration request to the AS, passing some information about the service (S_{info}) and the public key (K_{pub}^{S}) of the service: $S_{info} + K_{pub}^{S}$;

3. AS receives this information, verifies it and then creates a unique service identifier (S_{UUID}). After this verification the AS creates the service credentials that will identify this service globally and uniquely on the platform: $C_{S[UUID]}^{AS} = K_{priv}^{AS}\{S_{UUID}, K_{pub}^{S[UUID]}, C_{AS}^{AS}\}$[1]. These credentials, which are signed by AS, are then returned to the requesting service;

[1] Some notes about the notation used: key(content) means the "content" is encrypted using "key"; key{content} represents "content" is signed using "key"; algo[content] means that "content" is hashed with the "algo" algorithm.

4. The requesting service, stores the obtained credentials that will be used in the future. This credential contains also the public key of the authentication service (K_{pub}^{AS}). This is used to prove this credentials to other entities that also rely on the same AS – services that trust AS, also trust on credentials issued by AS, presented by other services.

The service registration process is repeated for the number of services available within the health information system. This enables the entire ecosystem of services to be trusted on that platform.

Another important registration process concerns the registration of the different actors that will interact with the HIS on the RMS. The registration of the actors on the RMS can be dependent or independent of the HIS. If the actors are directly registered by the HIS, it will be necessary to have a synchronisation between the HIS and the RMS. In this paper, it will be assumed that the actors registration will be handled by the RMS, and that the HIS will delegate this registration process on the RMS. This actor registration process is depicted in Figure 5 and described in detail in the following steps:

1. Assuming that the user/actor still has no account created, it starts the registration process on the HIS. The HIS redirects the user to the RMS AUTS, that will handle the user/actor registration on the HIS behalf;

2. The user/actor, making usage of the client-side RMS authorisation service (AUTS), initiates the registration process. AUTS presents different registration options to the end-user depending on some HIS registration requirements (this will depend on the specific scenario to be implemented by the HIS itself);

3. The user/actor enters all the HIS required registration information, including a set of traditional credentials (such as the email address and the password) on the AUTS;

4. The AUTS, using the user/actor selected credentials (email, password) creates a secret key (S_k) that is used to initialise a local secure storage (encrypted database) at the authorisation service: $S_k^{SStorage} = SHA1[email + password]$. This secure storage is used to securely store sensitive information at the end-user side, that will enable the information and content governance;

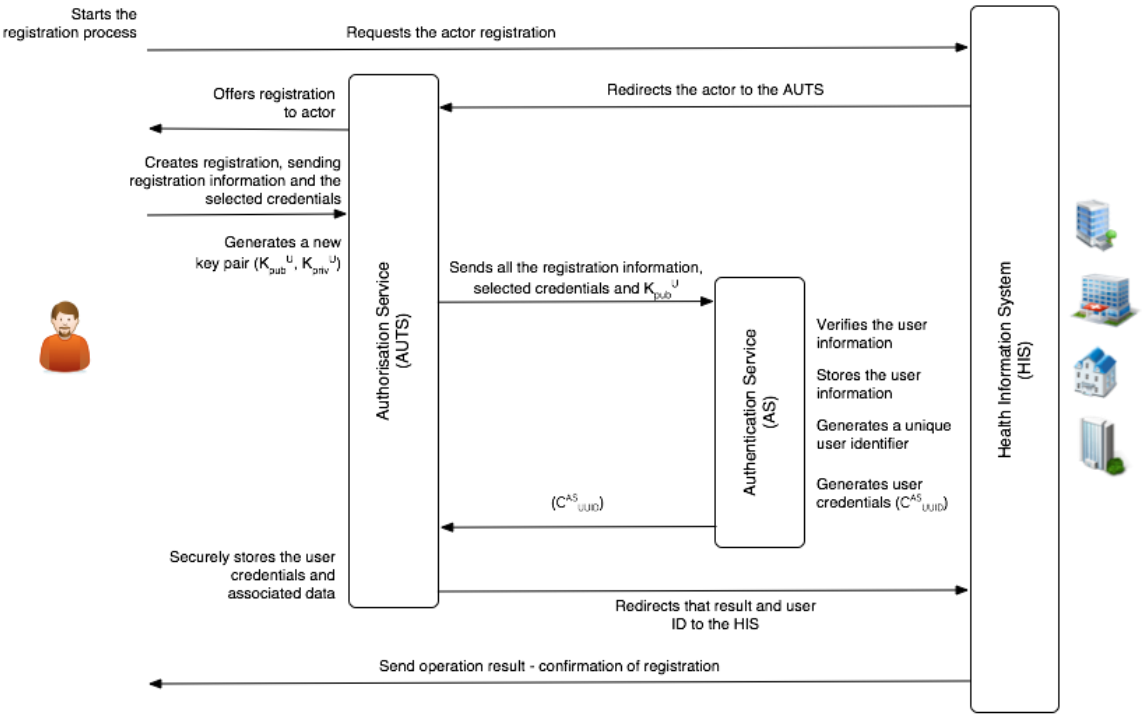

Figure 5. Process that describes the registration process on the system

6. The AUTS securely stores the user information. Additionally, the AUTS also creates a key-pair for the user/actor ($K_{pub}^{U}, K_{priv}^{U}$), storing it in a secure manner: $S_k^{SStorage}(K_{pub}^{U}, K_{priv}^{U}, user_{info})$;

7. AUTS contacts the AS to register the user on the RMS platform. This is performed using the $C_{S[AUTS]}^{AS}$ that contains the K_{pub}^{AS}. $C_{S[AUTS]}^{AS}$ is also sent to ensure that the AUTS has been previously registered: $K_{pub}^{AS}(email, K_{pub}^{U}, C_{S[AUTS]}^{AS})$;

8. The AUTS receives this information and after deciphering it, and validating the AUTS credential, registers the user/actor, generates a unique identifier for the user/actor (UUID) and creates credentials for that user/actor: $C_{UUID}^{AS} = K_{priv}^{AS}\{UUID, K_{pub}^{U}\}$;

9. The credentials are then returned to the AUTS and are securely stored: $S_k^{SStorage}(CA_{UUID}^{AS})$;

10. AUTS notifies the HIS about the result of the registration and sends the user AS credentials to the HIS (CA_{UUID}^{AS});

11. HIS notifies the user/actor about the result of the registration operation.

This represents the process that is used to register services and users/actors on the RMS platform. This will enable the creation of a trustworthy environment between the HIS, the RMS and all the different services that will be used to implement the information governance models, that will be responsible for maintaining the confidentiality and privacy of the health information.

In this process it will also be possible to ensure that there will be a decoupling between the HIS and the user/actors identification on the system, safeguarding the confidentiality of the user/actors personal information and preserving its privacy. The HIS will be able to identify users/actors simply by the UUID that was assigned by the RMS. Whenever more information needs to be disclosed by the user/actor, the user will have control over the information disclosed, when and to whom - controlled by the RMS.

The following section of this paper, introduces also an important characteristic of the system, that allows the secure storage of health-related information and associated content on the platform, and the definition of the appropriate information and content governance models.

CREATING AND PUBLISHING HEALTH-RELATED INFORMATION AND ASSOCIATED CONTENT ON THE PLATFORM

One of the objectives of the integration of the HIS with a RMS is the possibility to add information on the system, specifically about the PHR and EHR, that will be protected and governed by the RMS services. This will result in an information structure that will be protected and governed, allowing the control of its privacy (Figure 6).

This information structure will contain a set of different types of data, created by different entities and users/actors, that under the patient control, can define different information governance models, that could establish access models based on the identity, role, group, and operation. For instance, it will be possible to define a governance model where a nurse, with a given identity, belonging to the hospital X, will be able to read information about a given number of lab tests and results, for a certain period of time, about a given patient Y. At the same time, it will be possible to define another governance model where a doctor, with a given identity, belonging to hospital X, will be able to read information about

lab tests and results and historical medical data, add a specific medical report to the patient history, for a given period of time, about patient Y. Many other governance models can be defined.

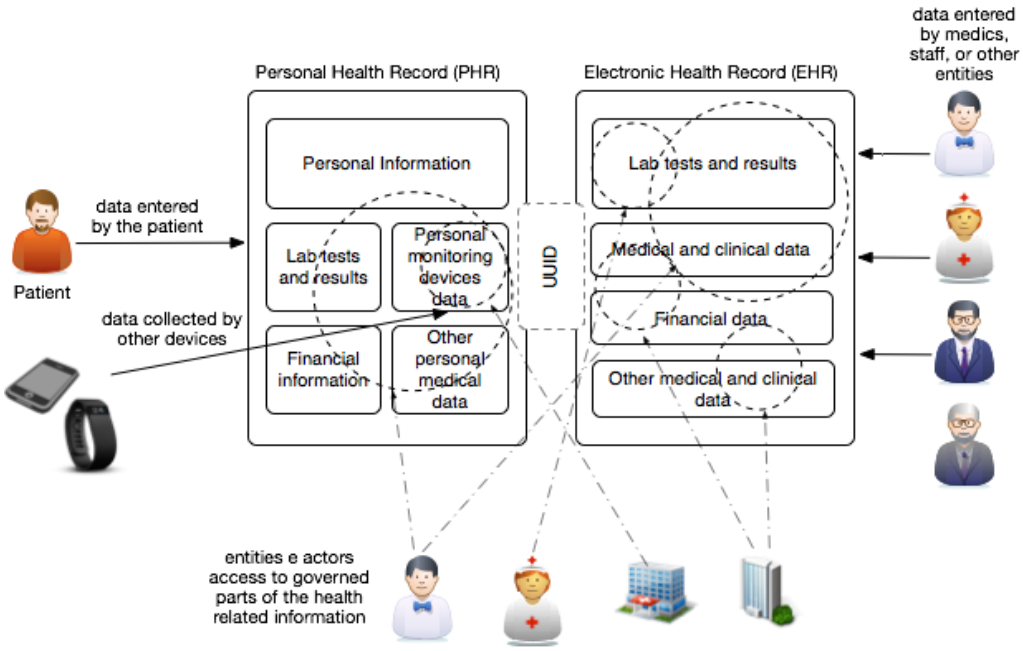

Figure 6. The scenario of the health-related information governance

Therefore the governance model can be defined by the following attributes:

- **User/Actor/Entity**: Unique identification of the user, actor, or entity. There will be the possibility to define groups of users and hierarchies between these different groups (for instance, it will be possible to say that doctor A and doctor B are part of the same medical team attending the patient C).
- **Affiliation**: If the user is part of an entity or a specific group, it will have to specify either the group identification (GIID) or the entity identification (EIID). Both the group and the entity can be defined in the RMS, and users/actors can be part of one or multiple groups or entities.
- **Rights**: Defines the different rights (operations) that are allowed over the health-related information and associated content - operations like, read, write, modify, and others.
- **Restrictions**: Defines the restrictions that are applied to the "Rights" that were assigned to the User/Actor/Entity over the Object (health-related information and associated content). There may exist different types of restrictions, such as temporal or geographic restrictions.

- **Delegation**: Defines either if this model of governance can be delegated to others or not, and what can be the length of the delegation path. If the model can be delegated, the Rights and Restrictions can be passed to third parties.
- **Object**: Contains the identification of the health-related information and associated content. The governance model will be specific for a given Object.

One of the most important goals of these governance models is to avoid that users, actors or entities have uncontrolled access to all the health-related data at any time. This will prevent data breaches that compromise the confidentiality and privacy of the data.

In order to ensure that this information governance models can be implemented the HIS health-related information and associated content needs to be published through the RMS. This, not only enables the governance model, but also at the same time allows health-related information and associated content to be stored securely on a configured location (for instance on the HIS cloud). Whenever a user/actor/entity/device creates and publishes new health-related information or related content, it is protected and the rights, permissions and restrictions about it can be defined by that user/actor/entity.

This creation and publishing process assumes that both user/actor/entity that generates the health-related information and associated content and the users/actors/entities willing to access the information are properly registered and authenticated on the HIS integrated with the RMS services (Figure 7).

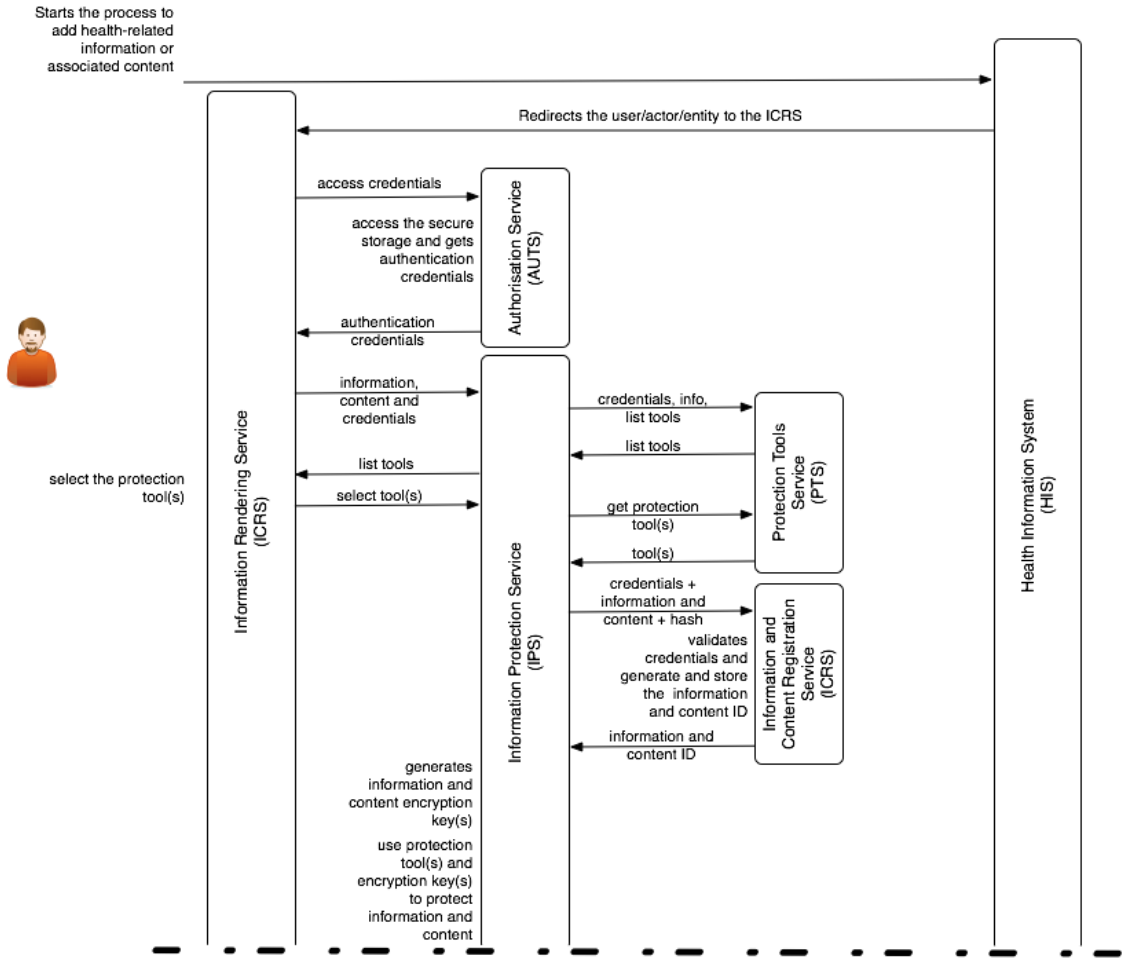

Figure 7. Adding and publishing information and associated content (Phase 1)

The overall objective of this process is to allow the creation and publishing of new health-related information and associated content the HIS governed by RMS services, define the access rights, permissions and restrictions, ensure that the content is protected, and return a location (URI) of the protected information and content to the HIS (Figure 8). This process starts with a request from a user/actor or entity to create and publish some information or content on HIS:

1. The user/actor/entity is redirect by the HIS to the Information Rendering Service (ICRS);
2. The user/actor/entity sends the health-related information and associated content (HIAC) that is relevant to be introduced on the HIS. This HIAC is uploaded through the ICRS. This service requires the user/actor/entity to enter its credentials (for instance, email and password or present any other form of previously registered authentication method), if it was not previously

authenticated. These credentials are used to access the secure storage: $S_k^{SStorage} = SHA1[email + password]$;

3. The ICRS contacts the AUTS, which reads from the secure storage the user/actor/entity RMS credentials: C_{UUID}^{AS};
4. The ICRS uploads the HIAC to the information and content protection service (IPS) and sends the user credentials, obtained in the previous step: $HIAC_{UUID}$, C_{UUID}^{AS};
5. The IPS, collects some metadata about the HIAC, contacts the protection tools service (PTS), requesting a list of available protection tools that can be suitable to protect the HIAC. The IPS sends its credentials and some information about the content: C_{IPS}^{AS}, $HIAC_{info}$. These tools will enable the adaptation of the RMS to multiple protection mechanisms, allowing the extensibility of the security tools used to offer protection to the HIAC;
6. The PTS returns a list of protection tools that match the request made by the IPS. This information is signed by PTS: $K_{priv}^{PTS}\{protection_{tools_{list}}\}$. This tool list will be different according to the type of HIAC and the characteristics of the content to protect;
7. The IPS returns the list of protection tools to the ICRS, and presents it to the user/actor/entity. They select the most appropriate protection tools, adjusting the parameters of applicability of the tools to the HIAC and submits its request about the necessary protection tools;
8. The IPS requests the selected protection tools from the protection tools service. The PTS returns the requested tools to the IPS;
9. Next, the IPS requests to the information and content registration service (ICRS) for the HIAC to be registered. For this, the IPS send its credentials, the HIAC metadata and the HIAC hash: $C_{IPS}^{AS}, HIAC_{info}, SHA1[HIAC]$;
10. The information and content registration service (ICRS), stores the received information, and generates a unique content identifier that is returned to the content protection service: $K_{priv}^{ICRS}\{HIAC_{UUID}\}$;

11. The IPS generates one or more HIAC encryption keys ($HEK_{[1][1]}, HEK_{[1][2]} \dots HEK_{[n][m]}$) that are applied over the HIAC, using the selected protection tools, in order to ensure the appropriate HIAC protection. There may exist multiple encryption keys for the multiple pieces of information and content that contain the HIAC;

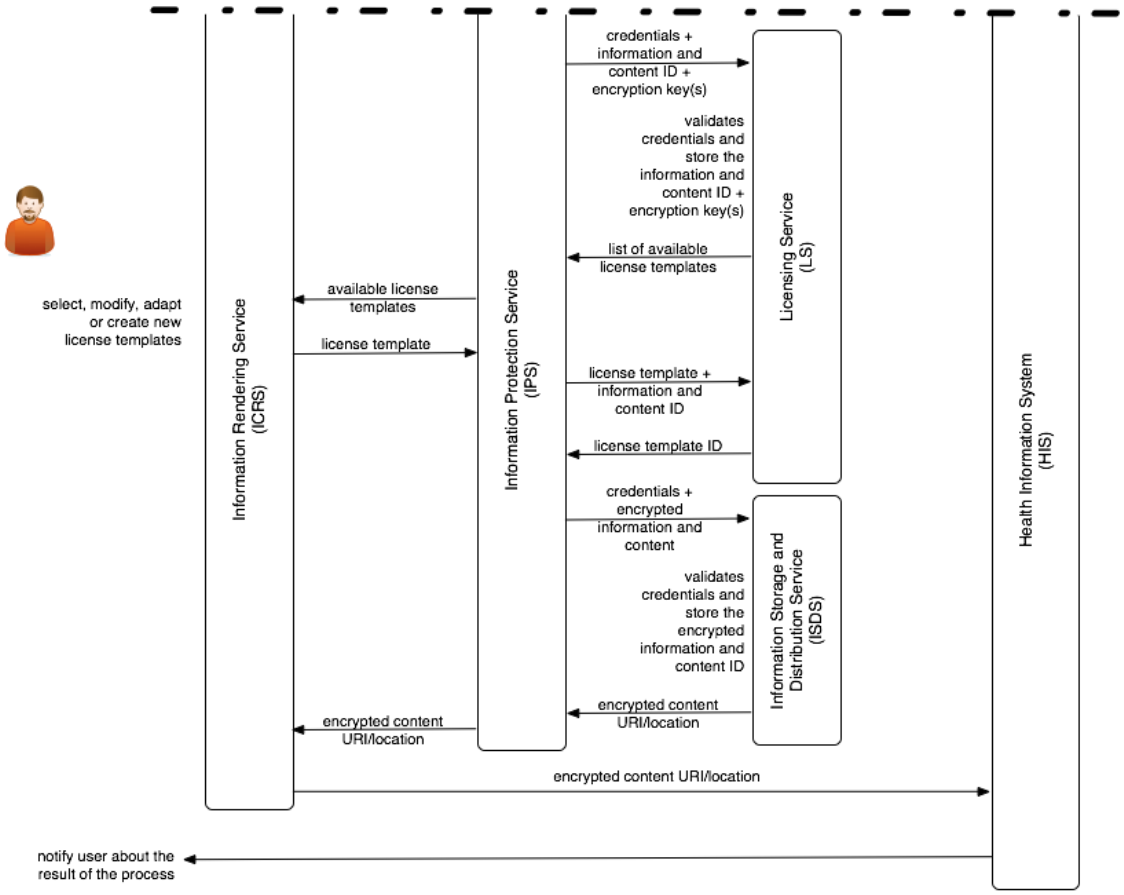

Figure 8. Adding and publishing information and associated content (Phase 2)

12. Following this protection process, the IPS sends the HIAC encryption keys for registration at the licensing service (LS). Each of the HIAC encryption keys is protected with the user/actor/entity key, and the entire message is protected by the IPS key: $C_{IPS}^{AS}, K_{pub}^{IPS}(K_{pub}^{U}(HEK_{[1][1]}, HEK_{[1][2]} \dots HEK_{[n][m]}), HIAC_{UUID})$;

13. The licensing service (LS) after validating all the received information returns a list of licensing templates to the IPS. The IPS returns the list of licensing templates to ICRS, and the user/actor/or entity can select the most appropriate license template, modify it and adapt it, or simply create a new one;

14. The license template (LIC_{TPL}) is sent to the IPS that after sends it to the LS and associates it with the identifier of the HIAC: $LIC_{TPL}, HIAC_{UUID}$. LS returns the license template identifier ($LIC_{TPL}[UUID]$);

15. In the next stage, the IPS sends the protected HIAC to the information storage and distribution service (ISDS) that stores the encrypted HIAC: $C_{IPS}^{AS}, K_{priv}^{IPS}\{HEK_{[n][m]}, HIAC_{UUID}\}$;

16. The content storage and distribution service returns a URI for the location of the stored encrypted HIAC. This URI is returned to the HIS that can reference afterwards.

After this process is completed, the HIAC will be store securely by the RMS services-enabled HIS. In order to have a fine-grained control over the HIAC, the user/actor/entity needs to use the RMS to produce specific licenses with the conditions under which the HIAC can be used. These licenses are produced in multiple formats (either in ODRL or MPEG-21 REL). In addition, these licenses are used to support the expression of rights over the HIAC and define the governance model. The following steps compose this process:

1. The IPS contacts the LS to obtain the appropriate license template for the specific HIAC, which was previously created: $LIC_{TPL}[UUID]$. The license template is an XML-formatted document that contains parameterized fields that can be adapted to specific rights situations - with the format that was also previously defined;

2. A typical license template for user generated content would be composed by following elements:

    a. User/Actor/Entity unique identifier (UUID), multiple users ($UUID_1, UUID_2, \dots UUID_n$) or a group identifier (G_{UUID}) (a group can be composed of multiple users, multiple entities or a combination of both): these fields represent the unique identifiers of the users or groups to whom the governance model will be established;

    b. The unique identifier of the content: $HIAC_{UUID}$;

c. List of permissions ($Permission_1..Permission_n$);

d. List of restrictions ($Restriction_1..Restriction_n$);

e. Validity date (validity);

f. The different HIAC encryption keys ($HEK_{[1][1]}, HEK_{[1][2]}, \dots HEK_{[n][m]}$). The HIAC encryption keys are protected with user/actor/entity public key: K_{pub}^{U} $(HEK_{[1][1]}, HEK_{[1][2]}, \dots HEK_{[n][m]})$;

g. The license signature, where the license contents are signed by the licensing service: $License = K_{priv}^{LIS}\{UUID_1..UUID_n,\ G_{UUID}1, G_{UUID}n,\ Permission_1..Permission_n,\ Restriction_1..Restriction_n,\ Validity, K_{pub}^{U}\ (HEK_{[1][1]}, HEK_{[1][2]}, \dots HEK_{[n][m]})\}$. In fact, this license, establishes the governance model.

3. The license is stored on the LS, where it can be accessed by legitimate users/actors/entities.

This concludes the entire process for creating and publishing HIAC on the RMS services-enable HIS. After this, all the entered information will be protected and governed, enabling the confidentiality and privacy of the EHR and PHR.

ACCESSING HEALTH-RELATED INFORMATION AND ASSOCIATED CONTENT ON THE PLATFORM

The last and definitive process to consider in this system is the access to information governed. The process that manages and controls who can access the HIAC, to which parts of the HIAC and what operations can be conducted is extremely important. In order for this to work, all the users/actors/entities need to be registered on the RMS services-enabled HIS.

The different users/actors/entities that need to access the HIAC, when navigating through the records and requesting operations over those records on the HIS, are “intercepted” and tunnelled through the RMS services and the governed access process is initiated:

1. The ICRS, while the user/actor/entity tries to conduct an operation on the protected HIAC, detects that it is governed, and contacts the AUTS to access the appropriate information to try conducting the operation on the HIAC;

2. The user/actor/entity authenticates to the system using the AUTS, supplying its credentials (for instance, email and password) to unlock the secure storage and retrieve its information;

3. The AUTS, using the $HIAC_{UUID}$ embedded on the HIAC URI, checks if a license for this HIAC already exists on the secure storage. If a license already exists:

    a. The AUTS checks the license contents, validating the license digital signature and verifying the $HIAC_{UUID}$;

    b. If the $HIAC_{UUID}$ is the right one, the Validity is checked and the list of permissions and restrictions are evaluated;

    c. If the conditions are met, the content can be deciphered and rendered by the ICRS and the operation requested can be conducted. The HIAC encryption keys can be retrieved from the license, and used to decipher the HIAC: $K_{priv}^{U}(K_{pub}^{U}(HEK_{[1][1]}, HEK_{[1][2]}, \dots, HEK_{[n][m]})) = HEK_{[1][1]}, HEK_{[1][2]}, \dots, HEK_{[n][m]}$;

    d. The operation over the HIAC can be authorised by the ICRS while the license conditions are fulfilled.

4. If the AUTS still does not possesses a valid license for the $HIAC_{UUID}$ that the ICRS is trying to conduct an operation over, the following steps need to be executed:

    a. The user/actor/entity authenticates to the system using the AUTS, supplying its credentials (for instance, the email and password) to unlock the secure storage and retrieve its information;

    b. AUTS, after getting the appropriated information, including the credentials, from the secure storage, allows the ICRS to contact the LS, passing its credentials (C_{ICRS}^{AS}), the

user credentials (C_{UUID}^{AS}) and the HIAC content identifier ($HIAC_{UUID}$) over which the operation is being request by a given user/actor/entity;

c. LIS receives and validates the data that was sent by the ICRS, and uses the HIAC content unique identifier ($HIAC_{UUID}$) and the user/actor/entity unique identifier (UUID) to verify the existence of a valid license. If the license exists on the system, that license is returned to the ICRS, that passes it, for validation and storage, to the AUTS: $License = K_{priv}^{LIS}\{UUID_1..UUID_n, G_{UUID}1, G_{UUID}n, HIAC_{UUID}, Permission_1 \dots Permission_n, Restriction_1 \dots Restriction_n, Validity, K_{pub}^{U}(HEK_{[1][1]}, HEK_{[1][2]}, \dots, HEK_{[n][m]})\}$;

d. The AUTS validates the license signature, verifying its contents and validity and asserting the correct $HIAC_{UUID}$;

e. If the $HIAC_{UUID}$ is the right one, the Validity is checked and the list of permissions and restrictions are evaluated;

f. If the conditions are met, the content can be deciphered and the operation requested can be conducted by the ICRS. The HIAC encryption keys can be retrieved from the license, and used to decipher the content: $K_{priv}^{U}\left(K_{pub}^{U}(HEK_{[1][1]}, HEK_{[1][2]}, \dots, HEK_{[n][m]})\right) = HEK_{[1][1]}, HEK_{[1][2]}, \dots, HEK_{[n][m]}$;

g. Operation over the HIAC can be authorised by the ICRS while the license conditions are satisfied.

In the conclusion of this process, the ICRS can authorise or not that the requested operation can be executed over the governed HIAC. After the operation is concluded, if the content needs to be re-protected, the system can take care of this as well.

CONCLUSIONS

The evolving technology has contributed to the digitalisation of several aspects of our lives. One of the areas that have embraced new information and telecommunication technologies is the health sector. Currently, information about our interactions with hospitals, clinics, doctors, hospital staff, exams and exam results, are all part of some databases on some remotely located datacentres. Although this represents an important step in the improvement of the healthcare services, allowing them to provide a better service to patients through the processing and analysis of large amounts of data, on the other side, the digitalisation of all this data represents a threat to confidentiality and privacy.

Cybercriminals are increasingly targeting this type of information, causing data breaches that provoke not only financial losses but also affect the reputation of entities, and expose in the wild private information that might lead to serious menaces such as identity theft attacks. Although there exist some legislative initiatives that intent to protect the privacy of health-related information, such as HIPAA, without the appropriate technological mechanisms, it will be impracticable to uphold such legislative principles.

The usage of rights management systems to offer confidentiality and privacy to health-related information and associated content, can offer a governed environment, that enables critical privacy and security mechanisms (Rodríguez, Rodríguez, Carreras, & Delgado, 2009). Health-related information can be governed by a rights management system to offer a finer control on privacy and security properties of EHR and PHR (T. Li & Slee, 2014). The proposed system on this article represents the attempt to demonstrate the applicability of an open and interoperable rights management system to enable information security and governance mechanisms that will improve the confidentiality and privacy of health-related information. The rights management solution enables the establishment of information governance models, through the usage of cryptography and rights expression languages, that will create the confidentiality and privacy environment for health stakeholders and users. This research proposes the usage of an open rights management system - OpenSDRM - based on a set of

decoupled services that can be integrated with any external Health Information Systems, to provide the necessary confidentiality and privacy requirements.

Throughout the paper it was possible to describe a set of core processes that enable the establishment of a trust environment between all the users/actors and entities, and between the different existing services. Also, the processes for the information and content publishing on the HIS, and the establishment of governance models, were defined. Moreover, this research also defined the processes necessary to enable the controlled access to governed health-related information and content.

Although we are conscious that our proposed approach and system will not solve all the health-related information confidentiality and privacy problems, at the same time we are convinced that it represents an important contribution towards the establishment of a decentralised governance model, that will reduce the impact of large data breaches, making harder for potential attackers the access to unprotected information. Since the work conducted is addressing in particular the prevention of non-authorised health-related data leakage, with a primary focus on confidentiality and privacy, independent from the protection system that is used to protect the data, or the governance model under which data can be used by different stakeholders, there are still some challenges that require further research. Considering the huge volume of health data that is captured from multiple sites and devices a particular interesting research direction consists in determining the overhead that using a rights management system can impact the timely access to data – in this particular case, and having into consideration that OpenSDRM can support multiple protection tools/mechanisms that can be applied to govern data, and assess the impact of the different ones on data access. A limitation from this work that also requires further research consists in the extension of the rights management system to enable the analysis of large amounts of governed health-related data while maintaining its confidentiality and privacy properties.

ACKNOWLEDGMENTS

This research has been conducted in the framework of the "Healthcare Insight – Units Performance Management" project (Ref[a] HI-UPM/2014/38567), and financed by the Regional Development

European Fund (FEDER), through the National Strategic Reference Frame (QREN), developed by the ISCTE-IUL School of Technology and Architecture.

REFERENCES

Abbas, A., & Khan, S. U. (2014). A review on the state-of-the-art privacy-preserving approaches in the e-health clouds. *IEEE Journal of Biomedical and Health Informatics*, *18*(4), 1431–1441.

AbuKhousa, E., Mohamed, N., & Al-Jaroodi, J. (2012). e-Health cloud: opportunities and challenges. *Future Internet*, *4*(3), 621–645.

Benson, T. (2012). *Principles of health interoperability HL7 and SNOMED*. Springer Science & Business Media.

Birnhack, M. D. (2008). The EU Data Protection Directive: An engine of a global regime. *Computer Law and Security Report*, *24*(6), 508–520. http://doi.org/10.1016/j.clsr.2008.09.001

Brailer, D. J. (2005). Interoperability: the key to the future health care system. *Health Affairs*, *24*, W5.

Brennan, P. F., Downs, S., & Casper, G. (2010). Project HealthDesign: rethinking the power and potential of personal health records. *Journal of Biomedical Informatics*, *43*(5 Suppl), S3–S5. http://doi.org/10.1016/j.jbi.2010.09.001

Buhse, W., & Van der Meer, J. (n.d.). The Open Mobile Alliance Digital Rights Management [Standards in a Nutshell]. *Signal Processing Magazine, IEEE*, *24*(1), 140–143.

Choo, K.-K. R. (2011). The cyber threat landscape: Challenges and future research directions. *Computers & Security*, *30*(8), 719–731.

Cushman, R., Froomkin, A. M., Cava, A., Abril, P., & Goodman, K. W. (2010). Ethical, legal and social issues for personal health records and applications. *Journal of Biomedical Informatics*, *43*(5 Suppl), S51–S55. http://doi.org/10.1016/j.jbi.2010.05.003

Djambazova, E., Almgren, M., Dimitrov, K., & Jonsson, E. (2011). Emerging and future cyber threats to critical systems. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6555 LNCS, pp. 29–46). http://doi.org/10.1007/978-3-642-19228-9_4

Fernando, J. I., & Dawson, L. L. (2009). The health information system security threat lifecycle: An informatics theory. *International Journal of Medical Informatics*, *78*(12), 815–826. http://doi.org/10.1016/j.ijmedinf.2009.08.006

Gerard Fernando, Tom Jacobs, & Vishy Swaminathan. (2005). Project DReaM, An architectural Overview. .

Haluza, D., & Jungwirth, D. (2015). ICT and the future of health care: aspects of health promotion. *International Journal of Medical Informatics*, *84*(1), 48–57.

Haux, R. (2006). Health information systems - Past, present, future. In *International Journal of Medical Informatics* (Vol. 75, pp. 268–281). http://doi.org/10.1016/j.ijmedinf.2005.08.002

Ikuomola, A. J., & Arowolo, O. O. (2014). Securing Patient Privacy in E-Health Cloud Using Homomorphic Encryption and Access Control. *International Journal of Computer Networks and Communications Security*, *2*(1), 15–21.

Jafari, M., Safavi-Naini, R., & Sheppard, N. P. (2011). A rights management approach to protection of privacy in a cloud of electronic health records. In *Proceedings of the 11th annual ACM workshop on Digital rights management* (pp. 23–30).

Juliadotter, N. V., & Choo, K.-K. R. (2015). Cloud attack and risk assessment taxonomy. *IEEE Cloud Computing*, *2*(1), 14–20.

Lazarovich, A. (2015). *Invisible Ink: blockchain for data privacy*. Massachusetts Institute of Technology.

Li, M., Yu, S., Zheng, Y., Ren, K., & Lou, W. (2013). Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *IEEE Transactions on Parallel and Distributed Systems*, *24*(1), 131–143.

Li, T., & Slee, T. (2014). The effects of information privacy concerns on digitizing personal health records. *Journal of the Association for Information Science and Technology*, *65*(8), 1541–1554. http://doi.org/10.1002/asi.23068

Liyanage, H., Krause, P., & de Lusignan, S. (2015). Using ontologies to improve semantic interoperability in health data. *Journal of Innovation in Health Informatics*, *22*(2), 309–315.

Marlin. (2006). Marlin Architecture Overview. .

Massey, A. K., Otto, P. N., & Antón, A. I. (2008). Aligning requirements with HIPAA in the iTrust system. In *Proceedings of the 16th IEEE International Requirements Engineering Conference, RE'08* (pp. 335–336). http://doi.org/10.1109/RE.2008.53

Meyer, J., & Pyles, J. (2005). The risks of healthcare IT. *Modern Healthcare*, *35*(31), 22. Retrieved from http://search.proquest.com/docview/211861700?accountid=14549\nhttp://hl5yy6xn2p.search.serialssolutions.com/?genre=article&sid=ProQ:&atitle=The+risks+of+healthcare+IT&title=Modern+Healthcare&issn=01607480&date=2005-08-01&volume=35&issue=31&spage=22&author

Nepal, S., Ranjan, R., & Choo, K.-K. R. (2015). Trustworthy processing of healthcare big data in hybrid clouds. *IEEE Cloud Computing*, *2*(2), 78–84.

OpenIPMP. (2003). OpenIPMP overview.

Redspin. (2014). *BREACH REPORT 2013: Protected Health Information (PHI)*. Retrieved from https://www.redspin.com/resources/whitepapers-datasheets/request-2013-breach-report-protected-health-information-phi-redspin.php

Rodrigues, J. J. P. C., de la Torre, I., Fernández, G., & López-Coronado, M. (2013). Analysis of the security and privacy requirements of cloud-based electronic health records systems. *Journal of Medical Internet Research*, *15*(8), e186.

Rodríguez, E., Rodríguez, V., Carreras, A., & Delgado, J. (2009). A Digital Rights Management approach to privacy in online social networks. In *Workshop on Privacy and Protection in Web-based Social Networks (within ICAIL'09), Barcelona*.

Saaranen, M., Parak, J., Honko, H., Aaltonen, T., & Korhonen, I. (2014). W2E—Wellness Warehouse Engine for semantic interoperability of consumer health data. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)* (pp.

350–354).

Serrão, C. (2005). Music-4you.com -- Digital Music E-Commerce Case Study. *IADIS International Journal on Internet/WWW*, *3*(1).

Serrão, C. (2008). *IDRM - Interoperable Digital Rights Management: Interoperability Mechanisms for Open Rights Management Platforms*. Universitat Politècnica de Catalunya. Retrieved from http://repositorio-iul.iscte.pt/handle/10071/1156

Serrão, C., Dias, J. M. S., & Kudumakis, P. (2005). From OPIMA to MPEG IPMP-X: A standard's history across R&D projects. *Signal Processing: Image Communication*, *20*(9), 972–994.

Serrão, C., & Dias, M. (2002). Space and Planetary Imaging using JPEG2000. In *Proceedings of the 7th International Workshop on Simulation for European Space Programmes*.

Serrão, C., Rodriguez, E., & Delgado, J. (2011). Approaching the rights management interoperability problem using intelligent brokerage mechanisms. *Computer Communications*, *34*(2), 129–139.

Serrão, C., Serra, A., Dias, M., & Delgado, J. (2006). Protection of MP3 Music Files Using Digital Rights Management and Symmetric Ciphering. In *Proceedings of the 2nd. International Conference of Automated Production of Cross Media Content for Multi-channel Distribution*.

Serrão, C., Serra, A., Fonseca, P., & Dias M. (2003). A Method for Protecting and Controlling Access to JPEG2000 Images. In *Proceedings of the International Society for Optical Engineering conference on Applications of Digital Image Processing XXVI (SPIE 2003) Anual Meeting* (Vol. 5203, pp. 272–286).

Sujansky, W. V, Faus, S. A., Stone, E., & Brennan, P. F. (2010). A method to implement fine-grained access control for personal health records through standard relational database queries. *Journal of Biomedical Informatics*, *43*(5 Suppl), S46–S50. http://doi.org/10.1016/j.jbi.2010.08.001

Sultan, N. (2014). Making use of cloud computing for healthcare provision: Opportunities and challenges. *International Journal of Information Management*, *34*(2), 177–184.

Thilakanathan, D., Chen, S., Nepal, S., Calvo, R., & Alem, L. (2014). A platform for secure monitoring and sharing of generic health data in the Cloud. *Future Generation Computer Systems*, *35*, 102–113.

Thomas, S. (2000). *SSL & TLS Essentials: Securing the Web* (Pap/Cdr). Wiley.

Torres, V., Delgado, J., & Llorente, S. (2006). An Implementation of a Trusted and Secure DRM Architecture. *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, 312–321.

Torres, V., Serrão, C., Dias, M. S., & Delgado, J. (2008). Open DRM and the Future of Media. *MultiMedia, IEEE*, *15*(2), 28–36.

Wang, B., Li, B., & Li, H. (2012). Oruta: Privacy-preserving public auditing for shared data in the cloud. In *Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on* (pp. 295–302).

Wu, R., Ahn, G.-J., & Hu, H. (2012). Towards HIPAA-compliant healthcare systems. In *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics* -

IHI '12 (p. 593). http://doi.org/10.1145/2110363.2110429

Zheng, Y.-L., Ding, X.-R., Poon, C. C. Y., Lo, B. P. L., Zhang, H., Zhou, X.-L., … Zhang, Y.-T. (2014). Unobtrusive sensing and wearable devices for health informatics. *IEEE Transactions on Biomedical Engineering*, *61*(5), 1538–1554.

Zhou, J., Lin, X., Dong, X., & Cao, Z. (2015). PSMPA: Patient Self-Controllable and Multi-Level Privacy-Preserving Cooperative Authentication in Distributedm-Healthcare Cloud Computing System. *IEEE Transactions on Parallel and Distributed Systems*, *26*(6), 1693–1703.