Article published online: 2022-12-09

FAIR Aspects of a Health Information Protection and Management System

Jaime Delgado[1] Silvia Llorente[1]

[1] Department of Computer Architecture, Universitat Politècnica de Catalunya, Barcelona, Spain

Methods Inf Med

Address for correspondence Jaime Delgado, PhD, Department of Computer Architecture, Universitat Politècnica de Catalunya, C/Jordi Girona, 1-3, Barcelona, ES 08034, Spain (e-mail: jaime.delgado@upc.edu).

Abstract

Background Privacy management is a key issue when dealing with storage and distribution of health information. However, FAIR (Findability, Accessibility, Interoperability, and Reusability) principles when sharing information are in increasing demand in several organizations, especially for information generated in public-funded research projects.

Objectives The two main objectives of the presented work are the definition of a secure and interoperable modular architecture to manage different kinds of medical content (xIPAMS [x, for Any kind of content, Information Protection And Management System] and HIPAMS [Health Information Protection And Management System]), and the application of FAIR principles to that architecture in such a way that privacy and security are compatible with FAIR.

Methods We propose the concept of xIPAMS as a modular architecture, following standards for interoperability, which defines mechanisms for privacy, protection, storage, search, and access to health-related information.

Results xIPAMS provides FAIR principles and preserves patient's privacy. For each module, we identify how FAIR principles apply.

Conclusions We have analyzed how xIPAMS, and in particular HIPAMS (Health content), support the FAIR principles focusing on security and privacy. We have identified the FAIR principles supported by the different xIPAMS modules, concluding that the four principles are supported. Our analysis has also considered a possible implementation based on the concept of DACS (Document Access and Communication System), a system storing medical documents in a private and secure way. In addition, we have analyzed security aspects of the FAIRification process and how they are provided by xIPAMS modules.

Keywords

- privacy
- security
- health information
- FAIR principles
- interoperability
- information protection and management system

Introduction

Since privacy is key when dealing with health information, there is a strong need to protect patients' medical documents from unauthorized access.

On the other hand, sharing health information for research purposes may help in detecting and preventing medical conditions both for specific patients and for the general population.

However, obtaining at the same time both features (i.e., privacy preservation and sharing) over medical information is

received
April 12, 2022
accepted after revision
September 19, 2022

DOI https://doi.org/10.1055/s-0042-1758765.
ISSN 0026-1270.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

a difficult trade-off. In this context, this article presents a system that aims at providing FAIR (Findable, Accessible, Interoperable, and Reusable) principles[1] for sharing patients' health information and, at the same time, preserving their privacy.

This system, xIPAMS (x, for Any kind of content, Information Protection And Management System), is able to manage different types of information: multimedia content (Multimedia Information Protection And Management System [MIPAMS]), genomic information (Genomic Information Protection And Management System [GIPAMS]), health records or documents (Health Information Protection and Management System [HIPAMS]), etc. This article describes xIPAMS and the particular case of HIPAMS (for health content).

xIPAMS is a modular architecture, following standards for interoperability, which defines mechanisms for privacy, protection, storage, search, and access to health-related information. We already described the concept of GIPAMS in Delgado and Llorente[2] and Llorente and Delgado.[3] However, the original idea behind xIPAMS system was defined for multimedia information in Llorente et al[4] and it was called MIPAMS. In that specific case, the focus was not privacy of information but access control to protected multimedia content. An important step forward in xIPAMS (and MIPAMS), to be introduced in this article, is that FAIR principles are now an explicit concern.

The structure of the article is presented in the next section on Objectives.

Objectives

Following FAIR principles is an objective when sharing data is a requirement. On the other hand, security and privacy are also a strong requirement when handling health information. The first aim of this article is to introduce a platform for managing health information in a secure and interoperable way. The second aim is to describe how privacy and security can be compatible with the provision of the FAIR principles. This is presented in the context of the introduced platform.

As stated later in the Background section, research is going on in relating FAIR principles with privacy. Our approach in this area is to provide privacy in a FAIR system by using a specific modular and interoperable architecture, which we introduce in the Methods section. This architecture is xIPAMS, a system able to manage any kind of digital health information in a secure and private way.

In the Results section, we show how xIPAMS supports FAIR principles and, at the same time, preserves patients' privacy, which is a major aim of our research.

Finally, some discussion and conclusions are provided on the difficulties and the cost to achieve our goal.

Background

The FAIR data principles consist of making data findable, accessible, interoperable, and reusable. They were first formally introduced in Wilkinson et al.[1] When data (very often "scientific data") are to be made publicly available, even subject to some conditions, a good approach is to achieve these principles.

When sharing medical information, to follow the FAIR principles may be an objective; in this case, provision of privacy should be considered. Some related works are briefly described next.

In Wilkinson et al,[1] the authors explain in detail the FAIR data principles. Then, they analyze how FAIR principles are provided in different organizations, including privacy considerations.

In Boeckhout et al,[5] the authors show the challenges and opportunities raised by the FAIR principles regarding medical data stewardship. One of the chapters in their study depicts how controlled data access provides privacy to the managed information.

In Wise et al,[6] the authors highlight that FAIR does not mean making the data open, as privacy and intellectual property issues may apply. In this case, access and authorization protocols may be required.

On the other hand, several projects are working, or have been working, on the definition of guidelines and on the implementation of tools following the FAIR principles, such as FAIRPlus[7] or FAIR4Health[8,9] projects.

The process by which data are converted or adapted to be FAIR is very often called FAIRification. This workflow consists of a set of steps that need to be followed to prepare the data. There are several initiatives defining those steps, although most of the approaches are very similar, as introduced next.

For example, GO FAIR,[10] an initiative that aims to implement the FAIR data principles, specifies its own FAIRification process. They propose guidelines to help making the data FAIR.

On the other hand, the FAIR4Health project[8,9] has developed its own workflow based on the FAIRification process adopted by GO FAIR and by taking into account the specific requirements identified in the health context, as further explained next.

The different steps of the FAIR4Health's FAIRification workflow could be summarized as:

1. Raw data analysis.
2. Data curation and validation.
3. Data de-identification/anonymization.
4. Semantic modeling.
5. Make data linkable.
6. License attribution.
7. Data versioning.
8. (Meta)data aggregation.
9. Archiving.

The most relevant steps to privacy and security are "Data de-identification/anonymization" (step 3) and "License attribution" (step 6).

On the other hand, the GO FAIR initiative also defines a step on licenses (called "Assign license"), making clear that "although license information is part of the metadata, they have incorporated the license assignment as a separate step in the FAIRification process to highlight its importance." It is very important to consider that in many situations having a license is the only way to access the data.

The Research Data Alliance (RDA)[11] is also very active in the area through several working groups (WGs), including CURE-FAIR WG, FAIR Data Maturity Model WG, FAIR for Research Software (FAIR4RS) WG, and Raising FAIRness in health data and Health Research Performing Organizations (HRPOs) WG. Furthermore, the RDA and FORCE11[12] have jointly created the FAIRsharing.org[13] registry of standards and other resources. The registry collects metadata to ensure that the information is FAIR, claiming that one way to achieve accessibility (the "A" from "FAIR") might be "by identifying their level of openness and/or license type."

Finally, in relation to the security and privacy aspects, GO FAIR refines the four principles, for example, with A1.2 (the protocol allows for an authentication and authorization where necessary) and R1.1 ((Meta)data are released with a clear and accessible data usage license). From this, the RDA identifies the importance of the evaluation of the fulfillment of these principles, what they called the "FAIR Data Maturity Model." Regarding the security and privacy-identified aspects, it means that data providers should evaluate if the access protocol supports authentication and authorization and if metadata refer to a standard license.

Methods

This section presents the xIPAMS, which defines an abstract, distributed interoperable architecture that deals with different kinds of digital content.

We have already defined similar architectures for multimedia information (MIPAMS[4]), and for genomic information (GIPAMS[2,3]). Based on the experience achieved when designing and developing them, we propose xIPAMS as an abstraction of the architecture, clearly defining common modules and trying to minimize the specific ones, which are left for the content-specific architectures. In this way, we can take benefit of the developments done for some kinds of content to provide support to other content types in an easier way, reusing modules as much as possible.

Any (x) Information Protection and Management System

As mentioned before, xIPAMS is an abstract architecture coming from the experience we have in defining secure and standards-based modular architectures for managing digital information and its associated metadata. The communication between the distributed modules is based on representational state transfer (REST),[14] which makes use of the HTTP protocol methods to communicate clients and servers.

As mentioned before, we defined MIPAMS,[4] for the management of multimedia information, and GIPAMS,[2,3] for the management of genomic information. In both cases, the modules that make up the architecture are based on well-known standards to favor interoperability. ►Fig. 1 shows xIPAMS modules and their relationships, together with the operations provided by each module at a high level.

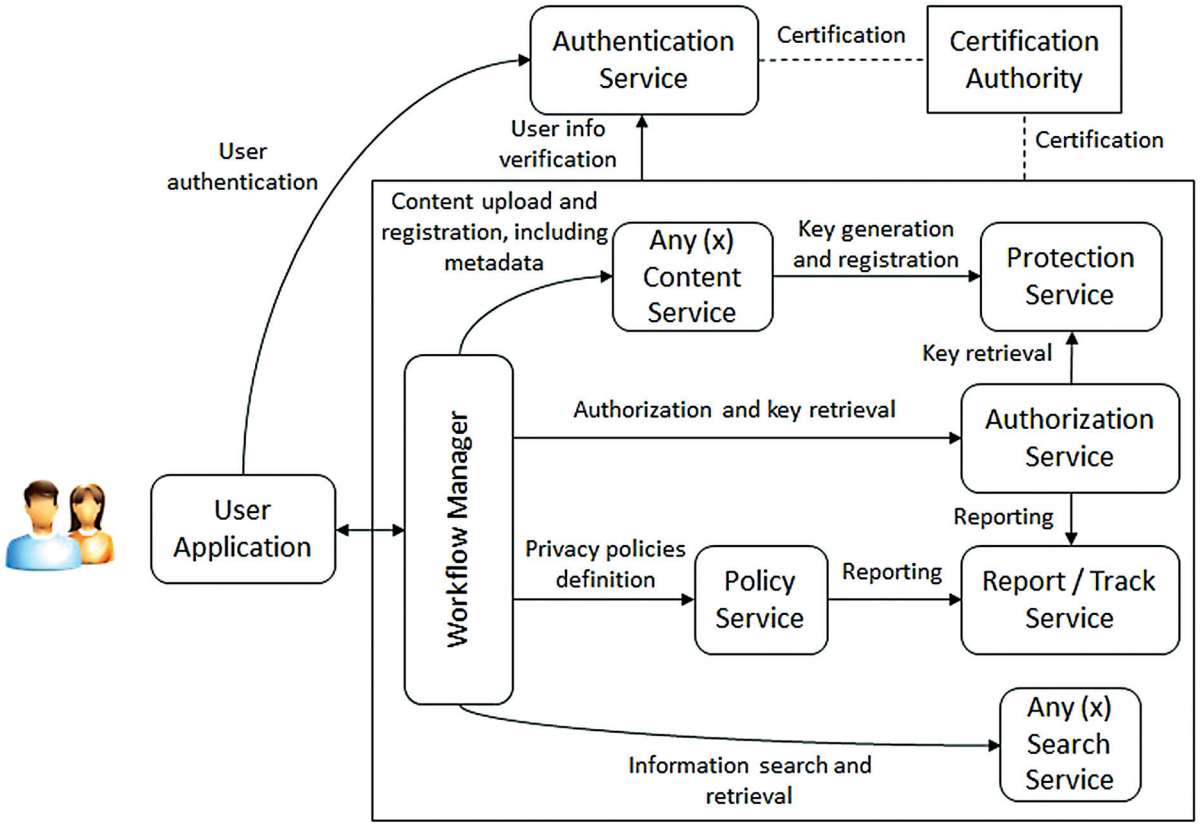

Fig. 1 xIPAMS architecture.

The functionality of the different modules is briefly outlined as follows:

- User Application (UA): an access point that sends all requests to the Workflow Manager, which redirects to the corresponding module based on the action requested by the user. An access token is required, which is provided by the Authentication Service. Communication between this application and the rest of the architecture is done through a secure channel. The UA may be implemented as a web application, a desktop application, or even a mobile application.
- Workflow Manager: an intermediate module that acts as a unique entry point to the system to facilitate interactions with the other modules, making them transparent to the final user. Before redirecting to the module in charge of an operation coming from the UA, it checks if this operation is authorized using the information inside the access token.
- Authentication Service: a server in charge of user identification. It provides authentication features using for example OAuth 2.0[15] and/or JSON Web tokens.[16]
- Content Service: a module in charge of digital content management, both in reading and writing operations. In case data have to be protected (encrypted and/or signed), it may connect to the Protection Service to get the required keys or the clear content. It also deals with metadata storage.
- Authorization Service: a module which validates authorization rules, for example expressed in eXtensible Access Control Markup Language (XACML).[17] Authorization requests are usually sent from the Workflow Manager responding to user actions, but other modules may also interact with it to request authorization of internal operations.
- Search Service: a module which performs searches over metadata associated to the specific digital content or even over the content itself. To provide extra filtering features, it may use a relational database where metadata fields are stored. It must be checked by means of authorization rules that the returned results may be shown to the user requesting them.
- Policy Service: a module in charge of the creation of the authorization rules. Again, it could make use of XACML[17] to create XACML policies and rules.
- Protection Service: a module that deals with keys and protection metadata, if any. It may also apply the defined mechanisms (i.e., encryption, signature, etc.) to the corresponding data or metadata.
- Report/Track Service: a module in charge of reporting the operations done in the system, especially those not authorized. It helps in keeping track of illegal/unusual operations that might indicate an attempt to attack the system.
- Certification Authority: this is not a real module of the system, but something required for its proper functioning. It provides the certificates needed to establish secure connections between the different system components.

xIPAMS Module Interaction

To show how xIPAMS modules interact among themselves, ►**Figs. 2** and **3** contain the sequence diagram of two system operations, Protected Content Creation and Authorized Content Search, respectively.

In ►**Fig. 2** we can see how the User interacts with the Authentication Service to get a token, which needs to be used when calling the different operations through the UA. After that, the different operations required to create and protect the content are invoked. Finally, some privacy policies are created.

In ►**Fig. 3**, we suppose that we have already obtained a token from the Authentication Service, so we directly invoke Search operation. Each search result obtained is authorized and only the authorized ones according to privacy rules are returned to the user.

Health Information Protection and Management System

As already introduced, the xIPAMS defines an abstract, distributed interoperable architecture that deals with different kinds of digital content. Therefore, the xIPAMS content-independent architecture may be then particularized for health information, resulting in HIPAMS (with "H" from "Health" content).

We should take into account that there are, in turn, different kinds of health information. We might have, for example, a HIPAMS managing clinical documents following specific standards. As an example for this case, we might consider a system to manage medical documents for both patients and medical staff, as in a DACS (Document Access and Communication System).[18] Documents could follow the Clinical Document Architecture (CDA) standard[19] from HL7 (Health Level Seven).[20]

A different situation, to be considered as future work, could be to combine in a HIPAMS system different kinds of eHealth content, such as the mentioned clinical documents, but also images and genomic information.

Results

As already stated, one of our objectives is to develop a system providing privacy and security following the FAIR principles.

One possible approach for this is to consider the concepts of Privacy by Design (PbD)[21] and Security by Design (SbD),[22] which are identified as the right approaches to reach the highest levels of privacy and security, respectively. In addition, we could consider a new concept: "FAIR by design" (FbD). Not to be confused with the idea of being "FAIR by design" (i.e., data are born FAIR) versus FAIRifying current datasets.[23] What we are considering here is the design of systems for data management and protection taking into account the FAIR principles. Developing the concept of FbD is future work based on the results of our current research.

Coming back to the PbD and SbD approaches, it is worth remembering that the European General Data Protection Regulation (GDPR)[24] in its article 32[25] mentions that technical measures for PbD and SbD must:

- Include pseudonymization and encryption of personal data (the GDPR does not explicitly deal with "anonymization").

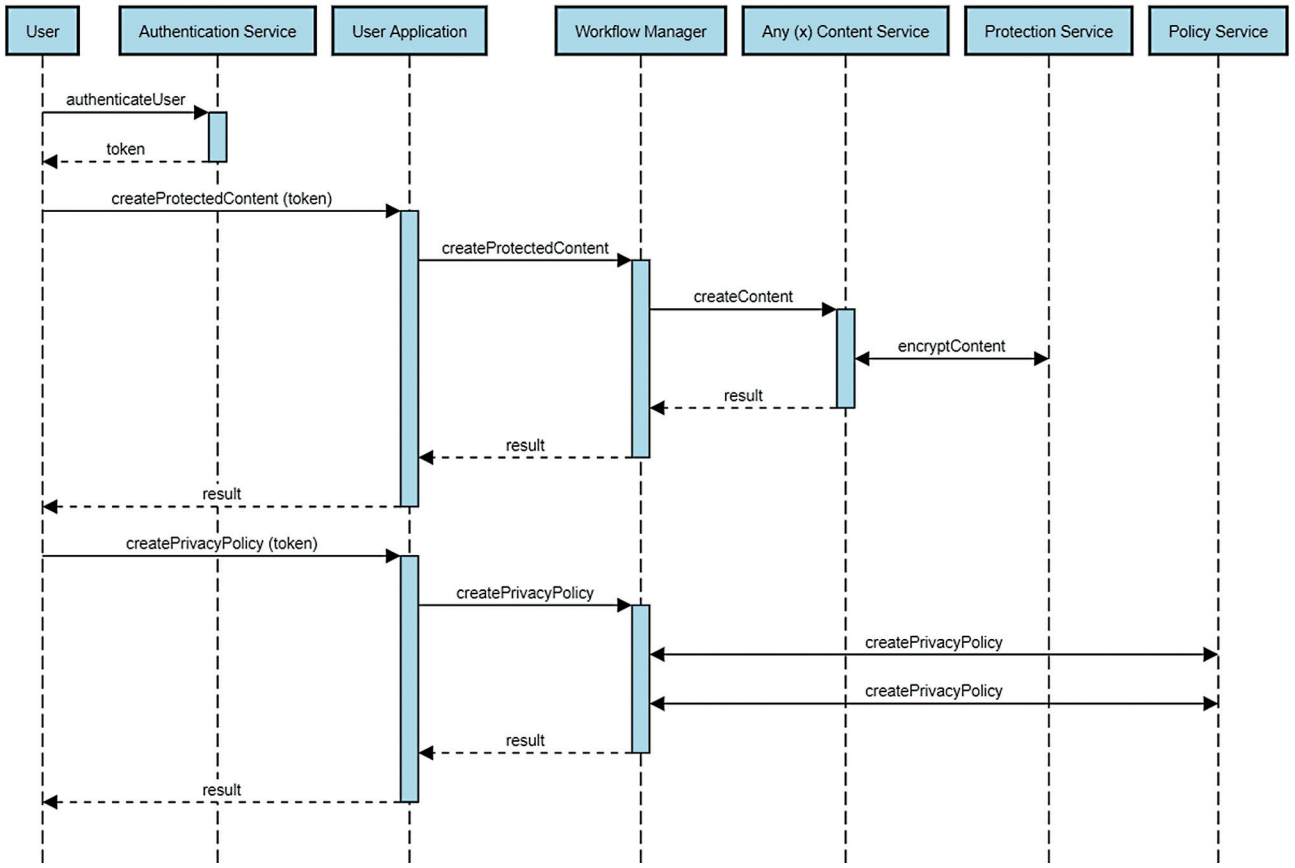

Fig. 2 Protected Content Creation sequence diagram.

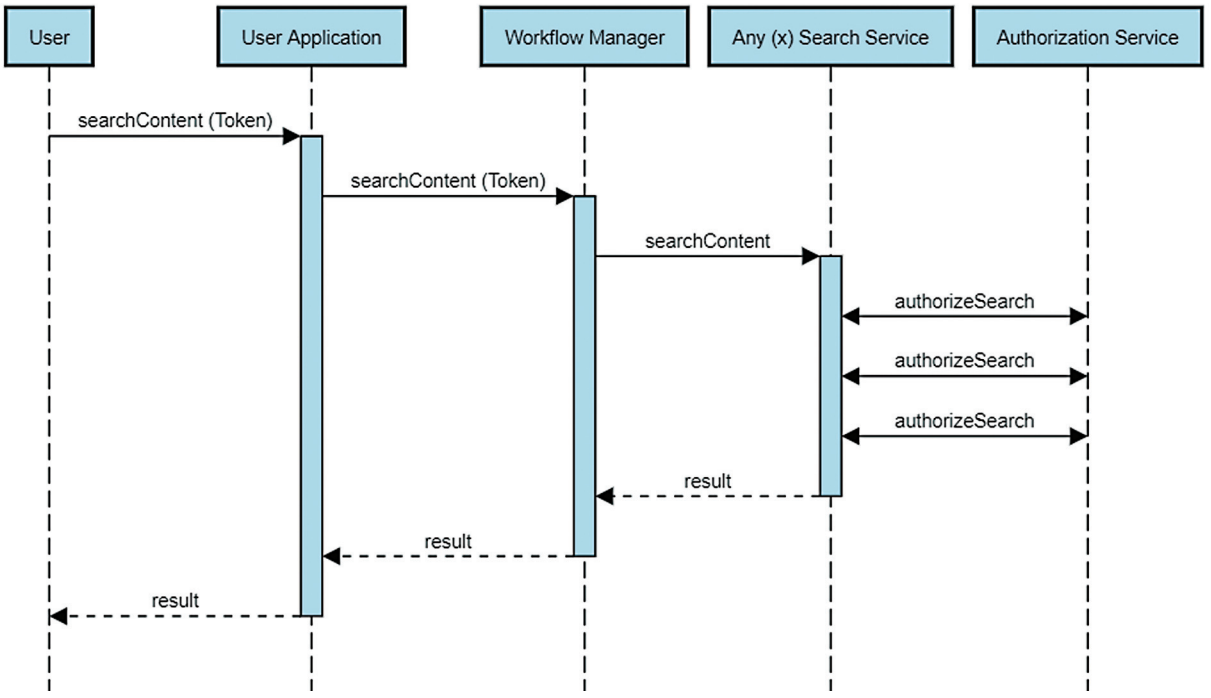

Fig. 3 Authorized Content Search sequence diagram.

- Ensure confidentiality, integrity, availability, and resilience of processing systems and services.

These two points are provided by xIPAMS, as already presented.

In addition, availability and access to personal data even after the event of a physical or technical incident must be provided. Finally, the GDPR requests for regular testing, assessing, and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.

In the context of the general approaches of PbD and SbD, these should be taken into account in all the steps of the FAIRification workflow introduced in Background section. It is of particular relevance for the steps where the information to FAIRify is obtained or updated, such as "1. Raw data analysis" and "2. Data curation and validation," or where security information is added, such as "3. Data de-identification/Anonymization" and "6. License attribution."

The design of xIPAMS followed the PbD and SbD methodologies and has taken into account GDPR requirements. This helps us in providing the required security and privacy for the FAIRification steps and to provide FAIR principles for the potential users of the information available in any xIPAMS platform or, in particular, a HIPAMS one.

In the rest of this section, we show how xIPAMS is compatible with FAIR principles and preserves patients' privacy.

The results presented in this section build on some of our previous work.[2,26] Delgado and Llorente[26] analyzed how to apply FAIR principles to genomic information considering privacy and security in the relevant FAIRification steps. It mainly considers steps 3 ("Data de-identification/anonymization") and 6 ("License attribution") of the FAIRification process described in GO FAIR.[10]

On the other hand, Delgado and Llorente[2] focus on GIPAMS, aiming at a first attempt to identify how it supports the FAIR principles. However, Delgado and Llorente[2] restrict the analysis to step 6; therefore, considering "license attribution" as the driving concept to analyze how the GIPAMS system is FAIR without losing privacy and security.

The approach we have taken in this section is twofold. First, to evaluate how the "License attribution" FAIRification step is supported and, second, to go through the four FAIR concepts and analyze how they are provided in xIPAMS.

License Attribution FAIRification step and xIPAMS

As introduced in the Background section, one of the main FAIRification steps with respect to privacy and security is License attribution. The objective of this step is to provide licensing attributions by data owners. We are considering access control rules (its expression and governance) in these FAIRification steps. There are many other options to express these licenses, but they come from the area of Digital Rights Management,[4] which is out of the scope of this article.

License attribution for datasets has to be clearly indicated. In addition, a process for a requester to ask for permission for getting "access" and therefore using, or in fact "reusing," a dataset is needed.

On the other hand, the absence of an explicit license may prevent others to access and reuse data, even if the data are intended to be open access.

Going deeper into the analysis of the License attribution concept in FAIR, in Delgado and Llorente,[26] we tried to answer four questions, namely:

1. How to express the licenses?
2. How to protect them and guarantee their provenance?
3. How to evaluate their authorization?
4. How to enforce what they are controlling?

We consider again these questions when handling licenses with xIPAMS. While Delgado and Llorente[2] just focused on GIPAMS, here we provide new answers to the four questions concerning licenses, from a FAIR point of view, but in an xIPAMS architecture.

Question 1: Expression of Licenses

The best way to express licenses is by using a formal language, thus facilitating interoperability (the "I" in FAIR). The policies and rules formally expressed define how to access the information (the "A" in FAIR) and allow for reusability (the "R" in FAIR). The standard we have chosen, and implemented in the Policy Service, to express licenses is the XACML.[17] The Policy Service is in charge of creating and storing the rules or licenses. XACML provides mechanisms to represent policies and rules that control who, how, and when may access specific information, be it data or metadata. The expression language is complemented with a mechanism to evaluate or validate the rules, based on standardized requests. While the Policy Service allows for the creation of rules, their evaluation takes place in the Authorization Service, which is the subject of Question 3.

Question 2: Protection of Provenance of Licenses

The Policy Service is also the one who answers the second question. In xIPAMS, the protection against modification of the created rules is achieved using XML Signature.[27] This is an extra feature for this service, as it also allows knowing the originator of the license. With respect to the FAIR analysis, this approach provides support to the Accessibility and even to the Interoperability, but also, partially, helps in providing Reusability, since we are confident in its origin and lack of modification. Therefore, this signature mechanism implemented in the Policy Service also provides the A, I, and R of FAIR.

Question 3: Authorization upon Licenses

The process of authorizing, or not, the access to specific content based on the licenses is provided in xIPAMS by the Authorization Service, which, in our implementation, uses the mechanisms defined in XACML.[17] For this purpose, we need to define "XACML Requests" including different attributes like subject, object, action, or time conditions. Then, the system evaluates if the attributes comply with any of the XACML rules available in xIPAMS. The request and the rule

are related to an object (data and/or metadata), which is stored in the Content Service module. Therefore, the A and I from FAIR are supported by the Policy and Authorization Services (as before), and also by the Content Service, to answer this question. Following with the Content module, we should notice that this module also provides for Reusability, thanks to the clear access rules associated to the content (data and metadata).

Question 4: Enforcement with Licenses

The enforcement assumes that the previous questions are solved, i.e., we have licenses and an interoperable mechanism to accept or not this access. If the requested action is authorized, nothing else special (except to get the content from the Content Service) is needed. On the contrary, we might need to protect the content, for example by encrypting it, to avoid unauthorized access. For this purpose, we have the Protection Service. Therefore, this service is regulating the Accessibility (A).

In addition, we have the possibility of keeping track of the actions performed in the system by means of the Report/Track Service, but this has no direct impact in the FAIR principles.

FAIR Principles and xIPAMS

The previous analysis has started from the License attribution FAIRification step as a guide and has identified how FAIR principles are implemented in each xIPAMS module for those purposes. This section, however, starts from the four FAIR principles and describes how the different xIPAMS modules provide every one of them. In addition, example implementations of HIPAMS are used to reinforce the fulfillment of FAIR concepts.

Findability

The only xIPAMS service not mentioned in relation to the analysis of the License attribution FAIRification step and xIPAMS (see the section "License Attribution FAIRification step and xIPAMS") is the Search Service. The added value of the Search module in the FAIR context is that it supports Findability (F), so with this we complete the support from xIPAMS to all FAIR concepts, since we had not assigned Findability to any of the xIPAMS services. Although the search function is not explicit for security, it is however relevant for FAIR.

The Search Service is queried through the UA, which should provide an easy interface for finding and retrieving content from the point of view of different kinds of users. Findable content could be medical documents, health records, genomic information, medical images, etc., depending on the specific xIPAMS.

If we focus on HIPAMS (Health content) and follow the example introduced in section "Health Information Protection and Management System," i.e., we consider a system to manage medical documents as in a DACS[18] with HL7[20] CDA documents,[19] those documents could be filtered using different parameters, for instance date, keyword, or signer, that is, which professional signed the document. Documents expressed using HL7 CDA are internally organized in sections, so, it is possible to recover only the relevant parts of them, instead of the whole document. For example, some of the sections that could be queried include Allergies, Medication, or Vital Signs.

Accessibility

Several xIPAMS modules help data to be Accessible, including the Content Service, together with the Policy, Protection, and Authorization Services, as already identified in the section "License Attribution FAIRification step and xIPAMS."

If we take again a DACS as an example implementation of xIPAMS (or HIPAMS, in this case), medical information can be presented in different ways: a complete document organized in sections or a composition of several documents whose link is the sections appearing in them. For instance, only the documents with a vaccines section may appear in the composition. Either the document or the document composition can be navigated in different ways, thus providing accessibility to the medical information they store. In the end, any structured medical document can be transformed to a PDF document, independently from the original document provided by the health professional, while storing the same information. This allows patients or other practitioners to access the documents without the need for a specific viewer, which is very often required in current health information systems (HISs), thus, improving accessibility.

Interoperability

As already introduced in the section "License Attribution FAIRification step and xIPAMS," Interoperability is provided by the Policy, Authorization, and Content modules.

Following up with our example, using a HL7-standardized CDA structure, instead of a proprietary one, facilitates interoperability between health organizations. A DACS provides medical documents that do not rely on a specific HIS' internal structure. In this way, moving documents from DACS to DACS (and then from HIS to HIS) is a feasible task, as they follow the same structure, which is the one provided by HL7 CDA. This applies to any standardized health information format. In any case, the use of standardized formats facilitates interoperability, as we could even make integration and adaptation (for example through conversions) between different formats of the same kind of content.

Reusability

The Reusability FAIR principle is implemented through a few xIPAMS modules, namely the Policy and Content modules, as stated in the section "License Attribution FAIRification step and xIPAMS."

It is always possible to reuse information stored in a HIPAMS implementation, that is, to use the information on stored documents with objectives (and of course users) different from the ones that the documents were intended for. For instance, a researcher would like to carry out an analysis over the effect of a specific treatment on patients with high blood pressure. In this case, the original documents were created to register a medical condition of the

patients, possibly storing the use of the treatment and if the patient suffered from high blood pressure. On the other hand, the researcher will only be interested in the documents explaining the effect of the medical treatment on the patient, if any.

In the previous use case, we achieve the reusability principle, as the medical information was originally intended for describing some medical conditions of the patient, and, later on, this information (or part of it) may be used for research related with this specific medical issue.

Summary of FAIR principles in xIPAMS

►Fig. 4 summarizes the FAIR principles implemented in every xIPAMS module.

The Policy and Content Services support three principles (A, I, and R), while the Authorization one supports A and I. Accessibility is also supported by the Protection Service. Finally, the Search Service is the only one that supports Findability.

Some Implementation Details in xIPAMS

We have already implemented some of the modules presented in this article. ►Fig. 5 shows some of the technologies used, as described next.

The Authentication Service is based on Keycloak,[28] an open-source identity and access management system. The Policy and Authorization Services are based on the XACML[17] language from OASIS[29] and implemented using WSO2 Balana.[30] The Content and Search Services make use of a relational database, which is mySQL.[31] For the certificate issuance, Let's encrypt[32] authorization authority is used. Most of the services are implemented using Java,[33] specifically using Java 2 Platform Enterprise Edition (J2EE),[34] both for implementing web applications and REST-based web services. We are using the Apache Tomcat server.[35]

Discussion

The use of a modular architecture based on standards for the representation, management, protection, and interchange of information facilitates the achievement of the FAIR principles. This is especially helpful for eHealth systems, where many different HISs co-exist.

Privacy and Security in eHealth Systems

A specific objective has been to provide at the same time FAIR principles and privacy. We have seen in the Results section that the different service modules in xIPAMS, a system providing privacy and security, allow offering the FAIR principles.

Nevertheless, including privacy and security inside eHealth systems implies some added difficulties, as the ones outlined below.

One difficulty is that health organizations or data owners have to define the access rules for their medical content. The concept of access control already exists in current HIS, but access control policies are usually not precise enough to be

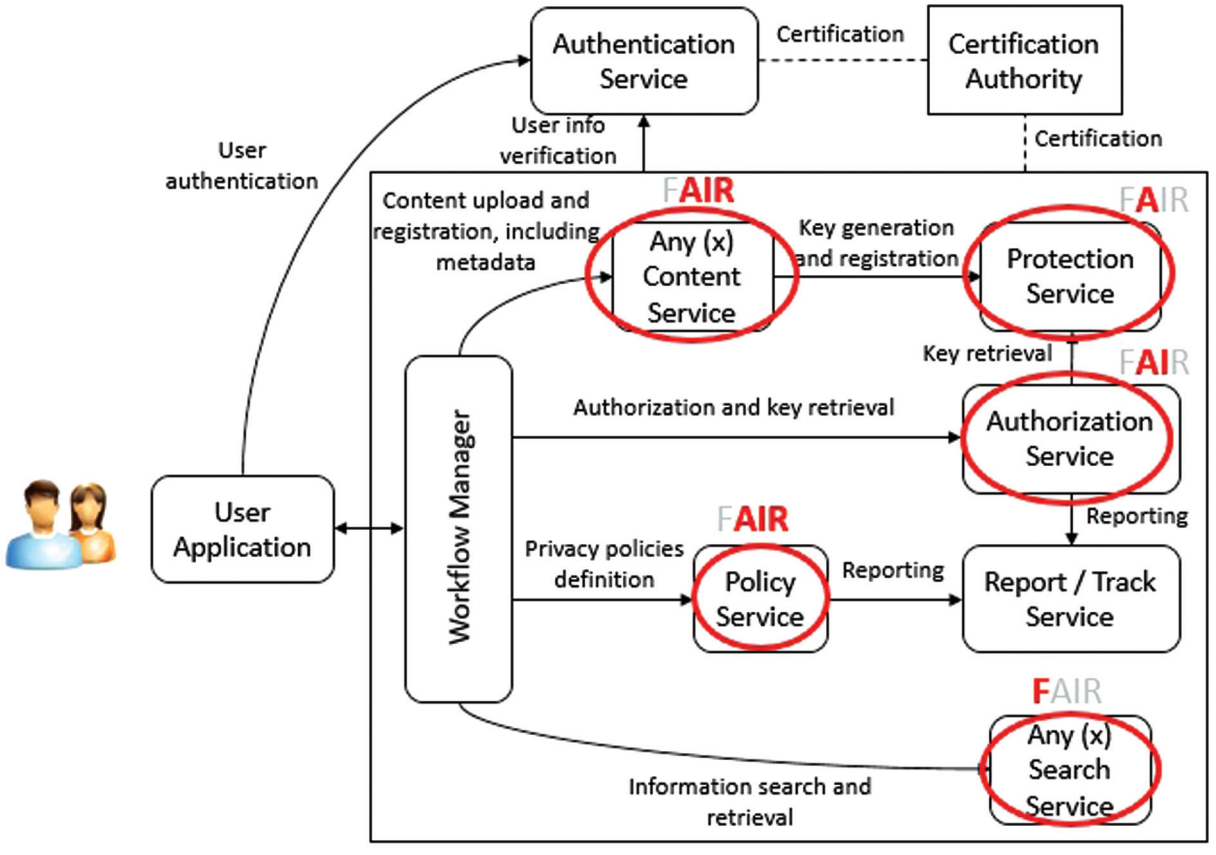

Fig. 4 FAIR principles in the xIPAMS architecture.

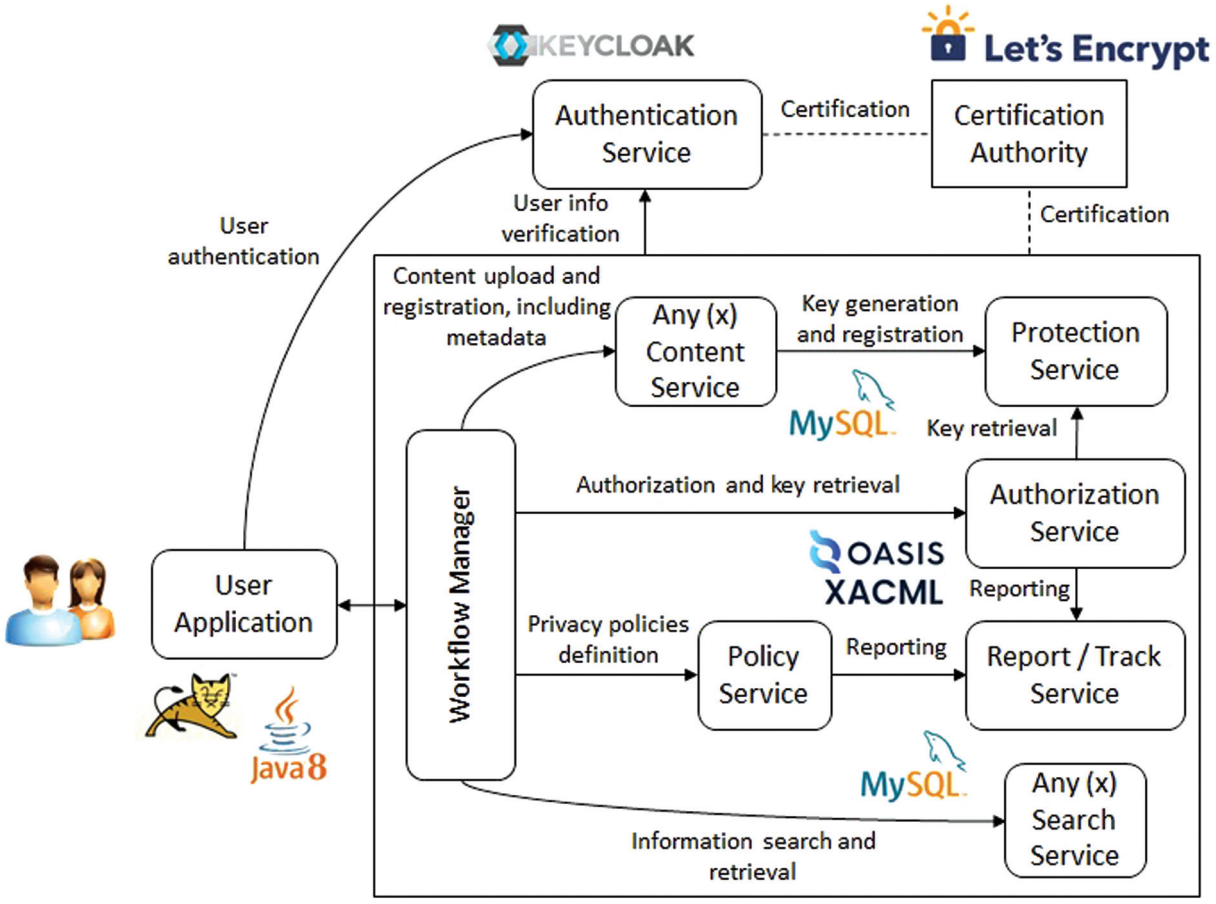

Fig. 5 xIPAMS architecture implementation details.

organized in formal access rules as the ones used in xIPAMS. Therefore, an effort to define a more complex access control model is necessary. In addition, it could be eventually an improvement of the security and privacy of the HIS.

On the other hand, the cost of accessing documents in a private and secure way increases, as new mechanisms have to be added to the system storing medical content. These mechanisms are those provided by the Authorization Service, for checking if access rules are accomplished, and the Protection Service, for the encryption of documents to prevent unauthorized accesses.

However, these difficulties and costs lead at the same time to important advantages, starting from the guarantee of security and privacy while implementing FAIR principles, that may also save relevant costs.

Interoperability of FAIR and Privacy and Security Information

In addition to the previous points, as we have previously shown, xIPAMS is interoperable as it relies on standards, and it is reusable as it provides requesting methods or operations which can be used to retrieve a subset of its content that is relevant in another context. Furthermore, the privacy rules in XACML allow providing an additional feature in the form of access control (which might be reinforced through the use of encryption).

It should be noted, however, that the XACML rules and the encryption mechanisms are not always integrated in the content to which they apply. This limitation happens for example in the case of CDA,[19] the standard from HL7 being used in many eHealth environments. In xIPAMS, this is clearly solved by the modular architecture, which makes content, licenses, and protection information independent. However, this might rise problems when moving data to another system. This is for example the case with this kind of HL7 documents, in which encryption parameters and XACML rules have to be transmitted in addition to the content.

A possible solution to this challenge is to define a container specification allowing the transport of the documents, encryption parameters, and privacy rules altogether. This would allow maintaining features even when moving the resources between repositories.

We have been also working in this kind of solution. In particular, we have a solution for the case of genomic information and in the context of ISO.[36] The standardization work has been done in one MPEG WG (ISO/IEC JTC1 SC29/WG8),[37] which has been working on standardizing a format for representing genomic data. The result of this work is the ISO/IEC 23092[38,39] standard, also known as MPEG-G (Genomic information representation). Part 3 of this standard[40] specifies how elements in a file should be encrypted, how the XACML privacy rules are to be included in the file, and how to

transport metadata relating to the content. It also defines a set of actions to be used as an interface with the content. The basic idea is the specification of metadata boxes, which are XML structures with the needed information integrated with the rest of the content.

Integrating this information inside the content still keeps the FAIR principles. Mechanisms similar to those introduced by ISO/IEC 23092 would allow enhancing each one of FAIR principles:

- Findability: the use of a standardized API (Application Programming Interface) simplifies the access to a subset of the content, allowing defining pointers to finer-grained objects.
- Accessibility: although it can be viewed as detrimental to a broad accessibility, the inclusion of the XACML rules within a standard file format allows defining a controlled and privacy-conscious access.
- Interoperable: having all this information structured within one container, whose structure is specified in a standard, ensures that migrating one file from one server to another is a simple process and does not imply a loss in features.
- Reusability: having metadata bundled with the data simplifies discovering what might be relevant in different situations.

Research Use versus Clinical Use

xIPAMS has been originally designed for secure content sharing. Therefore, in the context of health information, research (secondary use) is its main purpose. However, it is worth noting that xIPAMS can be also used for the clinical practice. To do so, we have to introduce the following specificities into xIPAMS modules:

- Content Service should work with clinical documents, like HL7 CDA.[19,20] In this way, xIPAMS may work like a DACS, considering clinical documents used in clinical practice.
- Policy Service should be able to define appropriate permissions for clinical use, like access for different user roles, such as nurse or doctor or specific specialties.

A new Metadata Service needs to be introduced to handle medical document-specific information to help in identifying and finding documents.

Conclusions

A modular, secure, and interoperable approach for the management of health information facilitates following the FAIR principles. We accomplish this with our xIPAMS, and in particular with HIPAMS (Health content). We have analyzed how xIPAMS supports the FAIR principles while keeping security and privacy. Our focus has been on rules to control the access to information. We have identified the FAIR principles supported by the different xIPAMS modules. The four principles are supported.

Implementing FAIR principles in a HIS is not an easy task due to the sensible nature of the information they store. To facilitate this task, our analysis has also considered a possible solution based on the concept of DACS, a system storing medical documents in a private and secure way.

In addition, we have analyzed security aspects of the FAIR-ification process and how they are provided by xIPAMS modules.

The combination of access rules, use of a standardized document structure, and addition of security mechanisms helps in accomplishing FAIR principles and preserving patient privacy. Privacy protection is key in our research and the approach presented here can be applied to other scenarios inside eHealth. This situation has been already identified and future work will combine in a HIPAMS system different kinds of eHealth content, such as clinical documents, images, and genomic information.

Another important topic that raised from this research is the new concept of "FbD." We will consider how to develop HISs with the aim of being FAIR, i.e., FbD. In addition, taking advantage of the experience of the xIPAMS specification, we will analyze how PbD, SbD, and the new FbD methodologies could work together.

On the other hand, alternatives to XACML as the standard to express access control rules and policies will be put in place.

Finally, we consider it worth noting that we plan to implement many of these solutions into a real clinical environment in the context of a new funded project in Spain.[41]

Ethical Considerations

Our manuscript does not involve research on human subjects.

Funding/Acknowledgments

The work presented in this article has been partially supported by the Spanish Government under the project: GenClinLab-Sec (Mechanisms for secure and efficient management of genomic information tailored to clinical laboratories: Security Aspects, PID2020-114394RB-C31) funded by MCIN/AEI/10.13039/501100011033 and by the Generalitat de Catalunya (2017 SGR 1749).

Conflict of Interest

None declared.

References

1 Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016;3:160018
2 Delgado J, Llorente S. FAIR aspects of a genomic information protection and management system. Stud Health Technol Inform 2021;287:50–54
3 Llorente S, Delgado J. Implementation of privacy and security for a genomic information system based on standards. J Pers Med 2022;12(06):915
4 Llorente S, Rodriguez E, Delgado J, Torres-Padrosa V. Standards-based architectures for content management. IEEE Multimedia 2013;20(04):62–72
5 Boeckhout M, Zielhuis GA, Bredenoord AL. The FAIR guiding principles for data stewardship: fair enough? Eur J Hum Genet 2018;26(07):931–936

6 Wise J, de Barron AG, Splendiani A, et al. Implementation and relevance of FAIR data principles in biopharmaceutical R&D. Drug Discov Today 2019;24(04):933–938
7 Harrow J. ELIXIR's Human Data Communities IMI FAIRplus project. Accessed July, 21, 2022 at: https://fairplus-project.eu/
8 FAIR4Health project. Accessed July, 21, 2022 at: https://www.fair4health.eu
9 Sinaci AA, Núñez-Benjumea FJ, Gencturk M, et al. From raw data to FAIR data: the FAIRification workflow for health research. Methods Inf Med 2020;59(S 01):e21–e32
10 GO FAIR. FAIRification process. Accessed July, 21, 2022 at: https://www.go-fair.org/fair-principles/fairification-process
11 Research Data Alliance Accessed July, 21, 2022 at: https://www.rd-alliance.org
12 FORCE11 (the Future of Research Communications and e-Scholarship). Accessed July, 21, 2022 at: https://www.force11.org
13 FAIRsharing.org Accessed July, 21, 2022 at: https://fairsharing.org
14 Fielding RT. Architectural Styles and the Design of Network-based Software Architectures [PhD dissertation]. University of California, Irvine; 2000
15 Internet Engineering Task Force (IETF) The OAuth 2.0 Authorization Framework. Accessed July, 21, 2022 at: https://datatracker.ietf.org/doc/html/rfc6749
16 Internet Engineering Task Force (IETF) JSON Web Token (JWT). Accessed July, 21, 2022 at: https://datatracker.ietf.org/doc/ html/rfc7519
17 Organization for the Advancement of Structured Information Systems (OASIS) eXtensible Access Control Markup Language (XACML) v3.0. Accessed July, 21, 2022 at: http://www.oasis-open.org/specs/index.php#xacmlv3.0
18 Delgado J, Llorente S, Pàmies M, Vilalta J. Security and privacy in a DACS. Stud Health Technol Inform 2016;228:122–126
19 ISO/HL7. ISO/HL7 27932:2009 Data Exchange Standards – HL7 Clinical Document Architecture, Release 2. Accessed November, 11, 2022, at https://www.iso.org/standard/44429.html
20 HL7 International Accessed July, 21, 2022 at: http://www.hl7.org/
21 Cavoukian A. Privacy by Design. Accessed July, 21, 2022 at: https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf
22 Open Web Application Security Project (OWASP) Security by design principles. Accessed July, 21, 2022 at: https://wiki.owasp.org/index.php/Security_by_Design_Principles
23 Singh S. The importance of a FAIR Data Strategy in enhancing the Data Value Lifecycle. Accessed July, 21, 2022 at: https://www.thehyve.nl/articles/fair-data-strategy-for-data-value-lifecycle
24 GDPR (General Data Protection Regulation) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC - Official Journal of the European Union, L 119, 4 May 2016. Accessed July, 21, 2022 at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL%3A2016%3A119%3ATOC
25 Article 32 of the GDPR (General Data Protection Regulation). Accessed July, 21, 2022 at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02016R0679-20160504
26 Delgado J, Llorente S. Security and privacy when applying FAIR Principles to genomic information. Stud Health Technol Inform 2020;275:37–41
27 World Wide Web Consortium. (W3C). XML Signature Syntax and Processing Version 2.0. Accessed July, 21, 2022 at: https://www.w3.org/TR/xmldsig-core2/
28 Keycloak. Open Source Identity and Access Management. Accessed July, 21, 2022 at: https://www.keycloak.org/
29 Organization for the Advancement of Structured Information Systems (OASIS) Accessed July, 21, 2022 at: https://www.oasis-open.org/
30 WSO2. WSO2 Balana. Accessed July, 21, 2022 at: https://github.com/wso2/balana
31 mySQL. mySQL relational database. Accessed July, 21, 2022 at: https://www.mysql.com/
32 Let's encrypt. Certificate authority. Accessed July, 21, 2022 at: https://letsencrypt.org/
33 Oracle. Oracle Java. Accessed July, 21, 2022 at: https://www.oracle.com/java/
34 Oracle. Java 2 Platform Enterprise Edition (J2EE). Accessed July, 21, 2022 at: https://www.oracle.com/java/technologies/appmodel.html
35 Apache Tomcat Accessed July, 21, 2022 at: https://tomcat.apache.org/
36 International Organization for Standardization (ISO) Accessed July, 21, 2022 at: https://www.iso.org
37 ISO/IEC. Moving Pictures Expert Group (MPEG). Accessed July, 21, 2022 at: https://www.mpeg.org
38 ISO/IEC. ISO/IEC 23092, Information technology—Genomic Information Representation. Accessed July, 21, 2022 at: https://www.mpeg.org/structure/genomic-coding/
39 Voges J, Hernaez M, Mattavelli M, Ostermann J. An introduction to MPEG-G: the first open ISO/IEC standard for the compression and exchange of genomic sequencing data. Proc IEEE 2021;109(09):1607–1622
40 ISO/IEC. ISO/IEC 23092–3, Information technology—Genomic information representation—Part 3: metadata and application programming interfaces (APIs), second edition. Accessed July, 21, 2022 at: https://www.iso.org/standard/82725.html
41 GenClinLab-Sec Project Mechanisms for secure and efficient management of genomic information tailored to clinical laboratories: Security Aspects, PID2020–114394RB-C31 funded by MCIN/AEI/10.13039/501100011033, Accessed July, 21, 2022 at: https://dmag.ac.upc.edu/proyecto_detalle.php?id_proyecto=50

