Security approaches for electronic health data handling through the Semantic Web: a scoping review

Vinícius Costa Lima[a,b,*], Filipe Andrade Bernardi[a,b], Domingos Alves[c], Rui Pedro Charters Lopes Rijo[d,e,f]

[a]*Health Intelligence Laboratory, Ribeirão Preto Medical School, University of São Paulo, Brazil*
[b]*Bioengineering Post Graduate Program, School of Engineering of São Carlos, University of São Paulo, Brazil*
[c]*Department of Social Medicine, Ribeirão Preto Medical School, University of São Paulo, Brazil*
[d]*School of Technology and Management, Polytechnic Institute of Leiria, Portugal*
[e]*Institute for Systems Engineering and Computers at Coimbra, Coimbra, Portugal*
[f]*Center for Health Technology and Services Research, Porto, Portugal*

Abstract. Integration of health information systems are crucial to advance the effective delivery of healthcare for individuals and communities across organizational boundaries. Semantic Web technologies may be used to connect, correlate, and integrate heterogeneous datasets spread over the internet. However, when working with sensitive data, such as health data, security mechanisms are needed. A scoping review of the literature was undertaken to provide a broad view of security mechanisms applied to, or along with, Semantic Web technologies that could allow its use with health data. Searches were conducted in the most relevant databases for the scope of this work. The findings were classified according to the main objective and features presented by each solution. Twenty-six studies were included in the review. They introduced mechanisms that addressed several security attributes, such as authentication, authorization, integrity, availability, confidentiality, privacy, and provenance. These mechanisms support access control frameworks, semantic and functional interoperability infrastructures, and privacy compliance solutions. The findings suggest that the application and use of Semantic Web technologies is still growing, with the healthcare area being particularly interested. The main security mechanisms for Semantic Web technologies, the key security attributes and properties, and the main gaps in the literature were identified, helping to understand the technical needs to mitigate the risks of handling personal health information over the Semantic Web. Also, this research has shown that complex and robust solutions are available to successfully address several security properties and features, depending on the context that the electronic health data is being managed.

Keywords: Semantic Web, Health Information Systems, Electronic Health Records, Computer Security, Interoperability

1. Introduction

In the current World Wide Web (WWW), most content is not easily accessible by machines, since it was made for human interpretation. The Semantic Web (SW) term was coined by Tim Berners-Lee and established by the World Wide Web Consortium (W3C), understood as an extension of the WWW that, besides of linking hypertext documents, can also recognize the information meaning and, through inference rules and ontologies, assist in knowledge management [6,59].

Integration of health information systems are crucial to advance the effective delivery of healthcare for individuals and communities across organizational boundaries [24]. In this sense, SW technologies can be used in open and sensitive contexts to connect, correlate, and integrate heterogeneous datasets spread over the Internet. However, when working with sensitive data, such as health data, security mechanisms are needed to protect it from unauthorized access.

Data leaks can cause harm in a variety of ways. A breach of security can result lives damage and in financial and legal consequences. For instance, improper handling of confidential data can violate government regulations, resulting in fines and other sanctions. Also, it may be a strong disincentive to data sharing initiatives among organizations [51]. In the case of health records, it usually represents the physical and/or mental condition of a patient. If security is breached, the disclosure of personal information may cause economic, social, and psychological potential harms [42].

The main goal of this scoping review is to provide a broad view of security mechanisms applied to, or along with, SW technologies that could allow its use with health data, as well as to identify possible research gaps in existing literature, and key characteristics or factors related to security in the SW.

It is expected to bring to the readers and the SW community an overview of security approaches in use and key related concepts. In this way, the thematic can be easily introduced to the general public, entry-level professionals, and interdisciplinary researchers, while helping advanced practitioners to quickly find solutions for their need.

The next section introduces the necessary background knowledge for the review. The third section explains the research methodology. The fourth section reports the review results. The fifth section presents a discussion based on the findings. Finally, conclusions are drawn in the last section.

2. Background

This section presents the definition of the assets and terms used as the basis for the review.

Health data is the asset being handled and that needs to be protected. This relates to any information about the physical or mental condition of a person, or to the provision of health services to the individual [17,32]. Health data can be stored in paper or in an electronic format, demanding distinct security measures. Paper-based information needs to be safely stored in controlled climatic conditions and protected against unauthorized physical access. In turn, a patient's digital health repository, i.e., the electronic health records (EHR), must be stored and exchanged securely, and accessible by multiple authorized users [21]. These requirements are essential to the digitization of health services towards a greater coverage of care, but are not trivial to achieve mainly due to the heterogeneity of the medical terminology and computational non-standardized technologies.

The SW plays an important role in the integration of disparate EHR systems due to the capability of expressing the meaning of a given information, i.e., its properties and the complex relationships between different types of data, in a way that enables the interpretation of its meaning without worrying about its form of representation [53].

However, EHR systems must comply with interoperability standards, i.e., they must be able to communicate with other systems in a transparent and consistent way [27] (functional interoperability), as well as to understand the context and meaning of the data provided by another system (semantic interoperability) [19].

Yet, when handling health data across entities, access control and privacy related concerns usually arise. The terms Confidentiality, Integrity, and Availability, also known as CIA triad, compose a base information security model to preserve the integrity and secrecy of stored or transmitted information [21].

Confidentiality considers that the data is only available to authorized personnel and, therefore, must not be disclosed to people who do not require them or who should not have access to them [21]. Integrity refers to the certainty that the data has not been subject to unauthorized deletion, modification or fabrication, either intentional or unintentional, during the storage or transmission [60]. Availability requests that an information must be available whenever necessary in a timely and uninterrupted manner [30].

Authentication, authorization and cryptography mechanisms are used to validate user's identity [44],

verify if the user has a sufficient level of access to perform an action (e.g., access control for read and write operations) [54], and encrypt data, respectively. They are essential components to guarantee the CIA triad and, therefore, provide a secure context to handle health information.

Finally, additional properties are desirable to increase the security level of a sensitive environment. Privacy control and provenance features allow system to comply with data protection regulations, e.g., the European General Data Protection Regulation (GDPR) [18]. Privacy is related to the right of no intrusion over one's personal information [20], avoiding the disclosure of any identifiable data. Provenance refers to the origination or source of specified data, such as requests to personal information, supporting privacy requirements and traceability of data flows [43].

Therefore, the CIA triad, authentication, authorization and cryptography mechanisms, and resources designed to manage privacy and provenance are critical to an effective security strategy for individuals' health information in the presence of deliberate or accidental threats and failures, as well as for risk management and information assurance practices.

Literature addresses security aspects related to the SW. Kirrane et. al. review the problems and solutions of privacy, security and policies with SW technologies [31]. Blanco et. al. presented a systematic review of security ontologies [7]. This work is focused only on ontologies and, therefore, leaves out solutions based on other SW technologies. There are specific reviews focused in the application of the SW in different areas, among others, Internet of Things [52], distance learning [4], and cloud computing [28], not evaluating security in depth.

3. Materials and Methods

Scoping reviews are ideal to provide an overview of a given topic, as well as to determine its coverage and give a clear indication of the volume of literature and studies available [41]. The methodology is based on Arksey and O'Malley [3] and guidelines provided by the Joanna Briggs Institute [47], which recommend a five-stage framework for scoping review. In the following subsections, each stage will be detailed.

3.1. Stage 1: Identifying the objective and research questions

This scoping review aims to verify the existing contributions in the literature to answer the following research questions:

Q1. Which are the main mechanisms being applied to, or along with, SW technologies to protect health data?

Q2. Which key security properties are being addressed?

Q3. What are the knowledge gaps regarding security for health data handling through the SW?

3.2. Stage 2: Identifying relevant studies

In this stage, the search strategy, i.e., the selected databases and keywords, and the eligibility criteria for assessing each primary study were defined to carry out the search to narrow the studies.

Search strategy. The following databases were searched: PubMed/MEDLINE, IEEE Xplore Digital Library, Scopus, Embase, Web of Science, ProQuest, and Cochrane Database of Systematic Reviews. These databases are of the most relevant in the scope of the present work, covering health and technology-related topics. The search string was defined as follows:

health AND "semantic web" AND (security OR privacy OR "access control" OR integrity OR confidentiality OR cryptography)

Eligibility criteria. Research papers, among others, full papers, reviews, and conference papers, and non-research studies, e.g., editorials, letters, in English language were included. The established time frame was from May 2001, when the term Semantic Web was first coined [6], to July 2021. Publications that do not refer to the keywords "*health*", "*semantic web*", and at least one security related term in the title or the abstract were excluded, as well as papers that present only proposals/non-implemented models or solutions not applied to the Semantic Web.

Forward and backward searching. The reference list of each selected article was searched to identify additional relevant studies that may have been left out in the database search. Eligibility criteria were applied to those relevant manuscripts that were included iteratively in the search, resulting in an accurate result.

3.3. Stage 3: Study Selection

Two investigators have independently screened each retrieved article based on title and abstract for

eligibility. Then, the full text was retrieved, and the investigators have performed another round of review. Fruitful discussions with the research team resolved the two reviewers' disagreements. Reviewers were not blinded to the journal's title, study authors, or associated institutions.

Although not being a systematic review, the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) [40] flow diagram was used to better comprehend the study selection.

3.4. Stage 4: Charting the data

An extraction strategy was defined to capture relevant data from the selected studies. Data extracted must be enough to answer the research questions established in Stage 1. Table 1 indicates the type of data that were extracted from each included article.

Table 1. Data extraction strategy

Scope	Data to be extracted
Summary	Title, authors, publication type, year of publication, periodic/journal, aims/objectives
Q1	Underlying mechanisms/technologies
Q2	Security properties/features
Q3	Benefits and limitations

3.5. Stage 5: Collating, summarizing and reporting the results

Since scoping studies seek to present an overview of all material reviewed [3], data were classified and presented in a table ordered by the category and by the year of publication. The table contains narrative content with data obtained in Stage 4. The narrative synthesis will seek to investigate similarities and differences between studies to explore patterns, themes, and relationships.

4. Results

Initially, 303 articles were found in the selected databases and 8 were identified through other sources (total of 311 articles), of which 197 were selected after removing 114 duplicates. After screening the titles and abstracts, the number of articles was reduced to 47. However, a full-text assessment for eligibility excluded 21 additional articles because 3 studies were inaccessible (no free or institutional access) and 18 did not meet the eligibility criteria. Finally, 26 (47-21) studies were included in the review. For simplicity, the diagram does not show all iterative cycles performed that included the addition of new manuscripts found in the analysis of all reference lists.

The Figure 1 shows the PRISMA flow diagram. Details of the selected studies are presented in Table 2.

According to the explored literature, research about security approaches to handle electronic health data through the Semantic Web has increased gradually since 2010. The results showed that the selected articles were published between 2005 and 2021, but the volume was higher between 2014 and 2019. Of the 26 included studies, there are 15 research articles, 10 conference papers and 1 editorial.

The following security attributes and features were addressed by the papers: authentication, authorization, integrity, availability, confidentiality, privacy, and provenance. Depending on the purpose of the solution, each study typically provides mechanisms to implement one or more of each attribute or features. Three categories were defined to classify the articles and reflect their objectives, namely Access Control (15 articles), Interoperability Infrastructure (3 articles), and Privacy Compliance (8 articles).

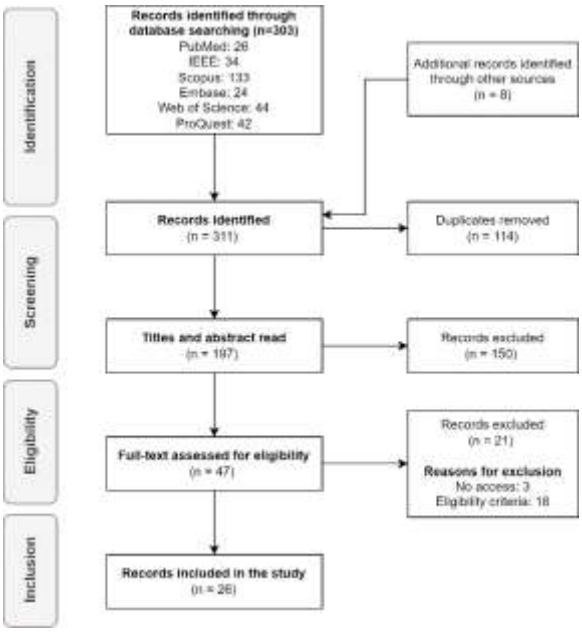

Fig 1. PRISMA flow diagram

4.1. Access Control

Studies in this category mainly present authentication and/or authorization mechanisms. However, other

features for confidentiality, privacy and provenance control are frequently available. All studies presented some authorization mechanism for health data integration through semantic web technologies [5,13,14,16,33–35,37,38,46,48–50,55,58]. However, only a few have implemented additional features for authentication [14,16,35,37,49,55], confidentiality [14,16], privacy [14,16,38,46,50], and provenance [46]. No studies in this category have addressed integrity and availability features.

Liu and Wang introduced a fine-grained context-aware access model for Health Care and Life Sciences (HCLS) Linked Data [37]. The authors use Semantic Web technologies to allow publishers of Linked Data to define access conditions for their data by extending the eXtensible Access Control Markup Language (XACML) with semantics. XACML rules define the policies and SWRL rules express semantic relations and inference problems. Automated decision making to permit or deny an access request is accomplished through inference processes based on the semantic relations among entities.

Li et. al. proposed a multi-layer authorization model that supports specifying and enforcing authorizations for pervasive healthcare delivery using Web Ontology Language (OWL) ontologies and semantic web technologies to conceptualize data and explicitly express the relationships among concepts and instances involved in information sharing [34]. The authorization model is composed of 4 layers, namely data/user, ontology, authorization, and application layers. Each layer handles a specific mechanism, such as data resources and organizations, concepts and relationships of objects, policies and rules, and the views, respectively. Authorizations can be specified at different levels of the predefined concept hierarchies and be propagated to lower-levels. Relying on ontology reasoning tools, the context dynamics must be encoded to enforce context-aware authorizations. Therefore, considering the source of the user's request, the data resource, the concept trees, the mapped relationships, the concept-level policies, and the security rules, semantic reasoning is conducted to obtain context-aware authorizations.

Rahmouni et. al. described a mathematical formalism for mapping SWRL privacy rules to standard access control based on XACML policies to avoid runtime overheads related to the enforcement of SWRL rules on complex and heterogeneous architectures [49]. The authors used Semantic Web technologies, such as OWL and SWRL, to model privacy requirements defined in European and national data protection laws as privacy aware access control policies. Through mathematical formalism, semi-automated mapping templates were established to transform the Semantic Web access control policies in XACML policies, a highly portable standard of access control. The solution could provide relatively easy interpretation of legislation at an operational level and the mapping of these ontologies to standard XACML policies could facilitate the implementation of SWRL rules in existing systems and complex environments.

Sun et. al. defined a Semantic based Access Control model (SAC) for e-Healthcare, which considers semantic relations among different entities in cloud computing environments, with the XACML standard to support description and management of distributed policies [55]. The SAC model extends role-based access control (RBAC) by considering the semantics of objects and associates permission with concepts instead of objects. The authors use ontologies for the RBAC security model and implement access control system in semantic web environment. An infrastructure was designed to enforce and evaluate authenticated requests. The SAC performs queries to the semantic knowledge base in order to find attributes associated with subjects and objects and translates the request to the XACML format. The XACML evaluates the request against an access control policy and sends the response to the requester.

Poulymenopoulou et.al. developed a security framework to provide discretionary role-based and context-aware access control services to read and update patients' data through a mobile application, which shares the user context in the requests sent to the cloud-based application server [48]. An OWL domain ontology was created to represent context information, enable context sharing and context reasoning. SWRL rules were written to capture the relationships between subjects and object. By using the ontologies and the SWRL rules (semantic knowledge base), the access control mechanism interprets the context and uses an inferring engine to grant access rights to patients' documents or to perform role changes for users (a user can move between permanent and temporary roles).

Yarmand et. al. proposed a flexible behavior-based access control for distributed healthcare systems [58]. The solution works by learning from the dynamic behavior of the user to determine access rights for clinical data. Ontologies are applied to map input factors from different environments to a standard format for semantic interoperability and, therefore, not requiring changes in the security architecture of each organization. The behavior of the user is compared with the expected user behavior and a decision-making engine

uses inference rules for reasoning to discover adequate privacy policies that should be applied for a user.

Dersingh et. al. propose a policy system to handle context and access management separately [13]. They extended the XACML language to demonstrate how contexts can be captured and represented semantically and integrated into an access control policy. The authors created an ontology to represent vocabularies (contexts) and compute context-based policies for access control. The contexts are acquired from a semantic knowledge base (a set of contexts/high-level domain vocabularies previously agreed) and the system suppresses user complexity of writing access control policies (rules) through the derivation of policies from the contexts. Finally, users' requests are evaluated by a proxy and confronted with the context-based policies (XACML request) to allow access to a given resource.

Lima et. al. introduced the implementation of authentication and authorization mechanisms in a semantic Application Programming Interface (API) - an API that produces responses in semantic formats, such as the JSON for Linking Data (JSON-LD) - based on access levels mapped to ontology properties, delivering granular access to a semantic tagged dataset [35]. This is a simple-but-robust mechanism to control which pieces of data a pre-authorized external system can retrieve from a semantic API. The solution allows granular access control to semantic annotated data, taking advantage of ontologies properties.

Kondylakis et. al. presents a novel access control mechanism to ensure the selective exposure of the patients' sensitive information [33]. The patient can provide electronic consent through a Personal Health Record (PHR) system for the use of him/her medical data system in different contexts. The patient monitors the access requests by the data consumers (e.g., doctors, nurses, insurance companies, etc.) and decides if data should be disclosed using a fine-grained access control. Instead of using standard annotation models, the solution relies on an abstract model to deal with the dynamics of the data and related accessibility. Data is associated with access labels (algebraic expressions and operators) and represented as a Resource Description Framework (RDF) data model to promote the interoperability among e-health systems and allow the protection of additional inferred information (along with explicit information). The algebraic expressions are interpreted in query time and, in events of consents' update or deletion, the costly redefinition of annotations for all triples in standard access control approaches is avoided. Complementary, Papakonstantinou et. al. demonstrate the use of the abstract model presented in [33] through the Health Access Control Enforcement Application (HACEA), where a user can disclose specific data (e.g., tumor type and malignity) for specific purposes (e.g., commercial, non-commercial, funding, emergency, and research) [46].

Dixit et. al. established a framework that enables policy based multi-authority access authorization to Electronic Health Records (HER) systems using the Multi-Authority Attribute Based Encryption (MA-ABE) and Semantic Web technologies to provide a semantically rich approach to facilitate secure data sharing among organizations who manage different attributes of end users using a shared dataset [14]. The solution implements a secure access control mechanism for user authentication and a robust crypto module for data encryption in order to tighten security and privacy before moving the data out of the organization. Attribute Based Access Control (ABAC) and a Multi-Authority EHR Ontology are used to carry out an access decision by matching the extracted attributes against the confidential access policies defined by an organization stored within the Policy unit in the form of SWRL rules. The use of a multi-authority ontology and SWRL rules provide flexibility for a complex environment where attributes differ among organizations. Therefore, reasoning is useful to evaluate an access request.

Rahmouni et. al. presented an ontology-based approach to tackle conflicting privacy and ethical requirements between national regulations in European countries by relying on Semantic Web technologies, such as SWRL rules for specification and reasoning of access control policies [50]. The authors suggest a direct mapping from high-level legislation on privacy and data protection to operational-level privacy-aware controls. In this case, OWL and SWRL are used for the specification and reasoning of access control policies, as well as user and data categories (as defined in legislations) and data disclosure contexts (e.g., maps the purposes of a needed consent). The paper defines an architecture for the enforcement of these controls on access control models adopted in healthgrid security infrastructures. A unique privacy context mode is involved when only one record of data is subject to a sharing request. A multi-privacy context mode is involved when large amounts of data need to be shared.

Lu and Sinnott presented a semantic based access control framework to extend the XACML with semantic capabilities to support fine-grained access control and ensure that privacy leakage can be detected and prevented [38]. The authors combine techniques of data anonymity, access control (XACML), and data semantics to show how privacy preservation can be satisfied through specifying background knowledge

and further restricting the access to certain data. According to the authors, the lack of semantic expressiveness is a barrier for finer-grained authorization. Therefore, the semantic framework includes policy formalization, compliance checking and knowledge discovery to prevent privacy risks with arbitrary linkages. To support semantic reasoning, policy vocabulary, domain knowledge and internal logic are mapped into ontological concepts. The mechanism of XACML systems is extended with Semantic Rules to achieve compliance checking between access requests and security policies. The proposed semantic framework has the potential to promote patient-centered healthcare, due to the possibility of accessing the complete historical information of the patient, including electronic health records stored in different health facilities.

Dridi et. al. developed a platform for the semantization of the Internet of Things (IoT) in the medical and healthcare field, regarding interoperability and integration of heterogeneous data, data visualization and access controls mechanisms [16]. The Semantic Medical IoT platform is capable of receiving data from medical equipment, IoT devices and electronic health records through a semantic interoperability layer, which performs semantic annotation and data integration. To ensure the security and confidentiality of information, the platform defines new contract-based security policies and provides a set of mechanisms for user's authentication and privacy control of personal health data. It is a comprehensive platform that deals with raw health data and transforms it into interpretable information (RDF format), while approaching security and privacy concerns.

Beimel and Peleg developed the SitBAC knowledge framework, a formal healthcare-oriented, context-based access-control framework able to represent patient's data-access scenario and perform inferences to either approve or deny access to data, based on OWL, a Description Logics (DL) reasoner and a SWRL engine [5]. The authors use ontologies defined in OWL to model Situation (scenarios) classes, formulating data-access rule classes. A set of data-access rule classes makes up the organization's data-access policy. The SWRL engine is used to infer new knowledge and relations. Then, the DL reasoner is used for knowledge classification and for real-time realization of the incoming data-access request as a member of an existing Situation class to infer the appropriate response. The inferred response type can be approved or denied. The OWL-based SitBAC knowledge framework complies with the “need to know” principle for data disclosure, which means that data is disclosed only when it is strictly necessary for someone to conduct official duties.

4.2. Interoperability Infrastructure

Studies have proposed more comprehensive architectures, capable of dealing with several of the security criteria mentioned above. These are generally more complex solutions, but they can support the functional and semantic interoperability of health data within a single framework.

Boniface and Wilken presented the ARTEMIS interoperability infrastructure for health information systems based on semantic web services and ontologies to broker between organizational policies [8]. The project delivers a complete architecture for functional, semantic and organizational interoperability, including security mechanisms for secure data exchange across organizations boundaries. Working as middleware, it can abstract the differences in security requirements (roles, clinical concepts and policies) and capabilities of each system through reasoning. The infrastructure enables the communication of standalone health information systems through mediation between semantic security and privacy policies, abstracting the differences in security requirements and capabilities of each system. The compatibility is achieved by using semantic web services and ontologies for reasoning of roles, clinical concepts and security policies.

Recently, Tiwari et. al. proposed the Secure Semantic Healthcare (S3HC) framework to represent, integrate and securely exchange data collected by healthcare devices [57]. It is a robust framework that delivers a semantic infrastructure for data storing, integration, and querying. The authors use an ontology designed for transferring the collected data from the device to the knowledge base and vice versa. A healthcare ontology named HCIoTO was designed to transfer the collected data from the device to the knowledge base and vice versa. SWRL rules and SPARQL queries are used to represent the accuracy and correct semantic reasoning between patients and doctors. Several security layers are available, such as RDF Security, XML Security and secure communication protocols. The framework collects, integrates and stores data from connected devices. To protect the data security properties such as confidentiality (encryption), integrity (hash functions), authentication (device identity verification), authorization (access policies), and availability are addressed. For patients, the main advantage is that the S3HC framework supports doctors

in analyzing the collected vital signs and in providing an appropriate and secure service for them.

Lima et.al. propose a framework for securing health data in a real case scenario focused on tuberculosis data exchange over the Semantic Web [36]. The solution is flexible and provides several endpoints (SPARQL, GraphQL and APIs) for functional and semantic interoperability of tuberculosis data. The framework is based on hybrid cryptography - a combination of symmetric and asymmetric techniques for encryption and transmission of big data -, hash functions and ontologies. It implements a Security Layer to support authentication and authorization for semantic web services and a query endpoint, ensuring confidentiality, integrity and availability of the data during exchange events. In addition, endpoints for SPARQL and GraphQL queries are available, enabling the extraction of tuberculosis health data from a regional health information system. The solution relies on ontologies and a virtual triple store database to convert, in real-time, legacy data stored in relational databases to semantic formats, so only responses in semantic formats (e.g., JSON-LD and RDF) are sent to authorized requesters.

4.3. Privacy Compliance

The concern with data confidentiality and personal privacy issues have been growing in the last years, driven by new regulations and laws. For patient health information, the Health Insurance Portability and Accountability Act (HIPAA) requires the creation of national standards to safeguard sensitive data from being disclosed without the patient's consent or knowledge [1]. Also, the GDPR aims to protect natural persons with regard to the processing of personal data [18]. SW technologies can underpin existing solutions to comply with these legal requirements in distinct contexts through their mapping and modelling.

Boussi Rahmouni et. al. presented an ontology-based approach for decision support regarding privacy in the sharing of patient data across European platforms through a SW application that can obtain privacy management guidelines for different entities in European countries involved in a data sharing process [9]. The authors use ontologies to model the required domain and context information about data sharing and privacy requirements and a set of Semantic Web Rule Language rules to reason about legal privacy requirements that are applicable to a specific context of data disclosure. A semantic web application is also available to provide decision support for clinicians to enhance privacy compliance. The application allows users to obtain privacy management guidelines for different entities in European countries involved in a data sharing process.

Alraja et.al. developed an integrated solution for users (data owners) of r applications to enhance privacy protection in events of private data sharing with a data consumer by calculating privacy risks associated with that specific sharing and comparing them to the benefits to-be received, providing a list of risks and recommendations to allow the user to take a pragmatic and informed decision [2]. The authors use an inference model, based on the Semantic Web and its supporting technologies (e.g., domain ontologies expressed using OWL) to allow the user to determine the privacy risks incurred when some personal data elements are shared with a data consumer. The framework provides useful information about the data request, such as the type of data being requested, the data consumer, and the context of the patient. Also, other data elements and accessible information about the same user are gathered (for example, from public data sources) and combined. Finally, a list of risks and recommendations is provided for the data owner. The feasibility and the utility of the solution is demonstrated by applying it to a case-study from healthcare and real patients. Based on the inference capabilities of the Semantic Web, the user can take an informed decision about the risks and benefits of sharing the personal data. It is a simple and powerful framework for privacy protection in IoT environments.

Joshi et. al. presented an OWL ontology to define the HIPAA stakeholders, as well as the privacy and security rules [29]. All concepts that have been specified in the act were defined, including business associates, covered entities (e.g., health care providers, health plans and clearing houses), personal health information, security safeguards, and administrative requirements. The ontology can be used by users to identify healthcare services that comply with the HIPAA regulations. Also, the ontology works as a guide to define security and privacy policies that should be implemented by the service provider to ensure HIPAA compliance.

Dong et. al. defined a Circle of Care (CoC) ontology that specifies concepts and relations necessary to capture a patient’s circle of care and allows one to make inferences about who is in a patient’s circle of care and, therefore, can access a patient's health records [15]. The proposed ontology and the easy integration with Health Level Seven International (HL7) FHIR supported EHRs is useful for explicit and implicit access consent. Access logs are annotated with

the COC ontology and converted into an RDF dataset that can be queried using SPARQL queries to investigate if an individual is in the circle of care of a patient. The solution help to overcome shortcomings of RBAC systems (e.g., capturing the consent of patients) and patient-centric approaches (e.g., patients may not be computer-savvy enough or have the necessary knowledge to be able to set permissions).

Relying on domain ontologies, Can and Yilmazer introduced a privacy-aware provenance management model to detect privacy violations and query provenance data, enabling traceability of historical data [10,11]. To preserve patients' privacy, the authors defined a healthcare provenance information system able to search for security violations based on access permissions defined by patients for their medical data. The solution relies on domain ontologies from different health fields to query, trace and protect sensitive data, as well as for the definition of access permissions.

Dao T.T. et.al. presented an approach to protect medical information using an asymmetric encryption algorithm with public and private keys, providing confidentiality for a semantic search engine to obtain Human Musculoskeletal System Resources (HMSR) information [12]. In their solution, Semantic web services process requests and a multi-agent crawler searches the World Wide Web (WWW) based on user-defined keywords. The search engine encrypts the results to protect medical information using a cryptography algorithm and a pair of keys. The user must use a private key to decrypt and read the search result.

Noor et. al. defined an Ontology for Detection of Attack on Genomic data (DAG) using semantic web technologies and a knowledge base of threats [45]. The system is able to analyze and validate incoming requests through inference rules. Incoming requests are parsed and potential attacks are compared with the information stored in the knowledge-base by inferring over the rules, and the system generates alerts upon detecting an attack. The ontology captures the context of attacks and threats on genomic data as well as potential consequences and the vulnerabilities exploited by these attacks for further analysis and to conduct mitigation actions.

Lastly, Celdrán et. al. proposed the h-MAS tool, a privacy-preserving multi context-aware solution, which allows users to choose profiles (e.g., privacy policies) to manage when, where, how, and to whom their private information can be revealed [25]. These profiles are specific to the context in which users are located, aimed to protect the privacy of their personal information, which can be modified by adding, modifying, or deleting its policies according to his/her interests. Information about users and contexts is represented by ontologies defined in OWL 2 and privacy policies are expressed in SWRL. Reasoning is performed to decide if a given information can be disclosed. The reasoner receives the ontological models generated by the Jena API [56] and applies SPARQL queries to obtain the requested information

5. Discussion

5.1. Overview of the different categories

In the articles in the Access Control category, although some works involve privacy control, this is done in the context of a broader mechanism that, through access policies based on rules and semantics, allows inferring whether a given agent has the necessary access rights, thus guaranteeing the privacy of individuals. The solutions usually work as a bridge between classical approaches for access control and SW technologies, but, usually, are complex to implement in existing health information systems.

On the other hand, articles in the Privacy Compliance category propose approaches focused on guaranteeing privacy by mapping requirements (e.g., laws and regulations) through models and ontologies, and they do not present a complete set of tools for access control.

In turn, the studies classified as Interoperability Infrastructure are more comprehensive as they provide tools that span the previous two categories through a complete infrastructure for functional and semantic health data interoperability.

In health, protecting data confidentiality is crucial. In Dao et.al [12], asymmetric encryption is performed, which deals with the problem of safely distributing the decryption key. Although the authors did not define access control mechanisms, they have demonstrated that using cryptography in sensitive contexts is an efficient way to ensure confidentiality and protect data from non-authorized readers, because only those in the possession of a decryption key are able to read the message. However, it is not clear how the authors deal with the amount of data that can be encrypted, due to limitations in the asymmetric cryptography [39], which may be a concerning factor to deal with big data.

Several studies share the same underlying technologies to deliver their solutions, including OWL/ontologies (all studies), SWRL [9,13,14,25,37,38,45,48–50], RDF [15,16,33,36,46,57], SPARQL [15,36,57], XACML [13,37,38,49,55], Jena [25,35,36,45],

Semantic APIs/Web Services [8,35,36], and Internet of Things (IoT) devices [2,16,57]. It demonstrates the flexibility of SW tools to allow the implementation of security mechanisms to protect sensitive data and still enable the interoperability and integration of such data. The semantic web may act as a bridge for the joint use of semantic technologies with classical ones, such as web services/APIs, XACML, and MA-ABE. Furthermore, it was observed that the use of these technologies in IoT devices is feasible to allow the traffic of sensitive data (e.g., personal health information), captured by these devices and safely transmitted over the internet.

5.2. Opportunities for future research

The findings of this work suggest that the main gaps in the literature refer to the absence of a complete computational architecture able to cover all the desired security properties in environments that handle sensitive data through the SW, and the high complexity of implementing existing solutions that, in most of the times, demand non-trivial changes in health information systems that were not initially developed considering the need of semantics.

Although the articles classified in the Interoperability Infrastructure category in this research may touch on those issues, they may represent a disincentive for the use of SW technologies to promote interoperability and semantic integration of health data. However, the development of a plug-and-play solution that could be offered as a service by trusted and secure cloud computing providers, as suggested by Alraja et.al [2], could be a way to simplify the implementation of security mechanisms in SW contexts.

5.3. General Recommendations

Ideally, confidentiality, privacy and provenance mechanisms should coexist in favor of an in-depth privacy compliance solution for health data handling. However, no study has presented a solution that considers all these mechanisms together.

Despite of not being designed for the Semantic Web, industry standards may reduce the impact of implementing security and semantic features. For instance, the Integrating the Healthcare Enterprise (IHE) nonprofit organization offers several profiles for security and privacy control (e.g. Document Encryption and Audit Trails) [26], while the HL7 FHIR defines exchange protocols and content models to be used with security protocols (e.g. Digital Signatures and Authentication) [22,23].

The research questions previously defined were satisfactorily answered. The main security mechanisms for SW technologies, the key security attributes and properties, and the main gaps in the literature were identified, helping to understand the technical needs to mitigate the risks of handling personal health information over the SW, and, ultimately, enabling the semantic interoperability and integration of such data.

6. Conclusions

In this research, the literature was explored to obtain a broad view of security approaches applied to the SW for electronic health data handling. The findings have shown that complex and robust solutions are available to successfully address several security properties and features, depending on the context that the electronic health data is being managed.

Although the Semantic Web paradigm was coined in 2001, the results suggest that the application and use of SW technologies is still growing, with the healthcare area being particularly interesting due to the SW inference capabilities for records linking and derivation of access policies. However, the complexity of a given solution tends to increase as more SW technologies and tools are incorporated, which can be seen as a disadvantage mainly for existing solutions not initially designed for the SW.

Even though this research was motivated by the healthcare scenario and involved the management of sensitive data, the SW is domain independent, since an adequate basis is available for its implementation (e.g., specific ontologies). The publishing of open linked data through the SW usually does not demand complex security mechanisms, but the security approaches identified in this study may be adapted to other scenarios that intend to use SW with sensitive data.

Finally, the SW community, academics, and beginners and advanced web professionals may take advantage of the findings of this work as an introductory material or to endorse the use of SW technologies in the health area.

6.1. Limitations of the study

This study presents some limitations. Although a satisfactory number of articles were selected, the inclusion and exclusion criteria restricted the scope and the possible applicability of Semantic Web

technologies and the associated security mechanisms, due to the interest of seeking for solutions applied into the health field. Therefore, the search was not exhaustive and future work can include research conducted in other areas, as well as to carry out comparison between them.

According to the methodology, articles that do not include the keyword "semantic web" in the title or the abstract were excluded. In some cases, this term might be missing or masked. To mitigate the risk of excluding important studies, the reference list of each selected article was searched to identify potential studies for inclusion.

Lastly, no quality appraisal was performed.

Acknowledgements

This study was supported by the São Paulo Research Foundation (FAPESP) [grant number 2020/01975-9].

Disclosure of any conflict of interest

The authors declare no conflicts of interests.

References

[1] 104th United States Congress, T., HIPAA. Health Insurance Portability and Accountability Act of 1996., 1996.

[2] Alraja, M.N., H. Barhamgi, A. Rattrout, and M. Barhamgi, An integrated framework for privacy protection in IoT — Applied to smart healthcare, *Comput. Electr. Eng.* **91** (2021) 107060. doi:10.1016/j.compeleceng.2021.107060.

[3] Arksey, H., and L. O'Malley, Scoping studies: Towards a methodological framework, *Int. J. Soc. Res. Methodol. Theory Pract.* **8** (2005) 19–32. doi:10.1080/1364557032000119616.

[4] Bashir, F., and N.F. Warraich, Systematic literature review of Semantic Web for distance learning, *Interact. Learn. Environ.* (2020). doi:10.1080/10494820.2020.1799023.

[5] Beimel, D., and M. Peleg, Using OWL and SWRL to represent and reason with situation-based access control policies, *Data Knowl. Eng.* **70** (2011) 596–615. doi:10.1016/j.datak.2011.03.006.

[6] Berners-Lee, T., J. Hendler, and O. Lassila, The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities, *Sci. Am.* **284** (2001) 34–43. doi:10.1038/scientificamerican0501-34.

[7] Blanco, C., J. Lasheras, R. Valencia-García, E. Fernández-Medina, A. Toval, and M. Piattini, A systematic review and comparison of security ontologies, *ARES 2008 - 3rd Int. Conf. Availability, Secur. Reliab. Proc.* **1** (2008) 813–820. doi:10.1109/ARES.2008.33.

[8] Boniface, M., and P. Wilken, ARTEMIS: Towards a secure interoperability infrastructure for healthcare information systems, *Stud. Health Technol. Inform.* **112** (2005) 181–189.

[9] Boussi Rahmouni, H., T. Solomonides, M. Casassa Mont, S. Shiu, and M. Rahmouni, A model-driven privacy compliance decision support for medical data sharing in Europe, *Methods Inf. Med.* **50** (2011) 326–336. doi:10.3414/ME10-01-0075.

[10] Can, O., and D. Yilmazer, Improving privacy in health care with an ontology-based provenance management system, *Expert Syst.* **37** (2020) 1–18. doi:10.1111/exsy.12427.

[11] Can, O., and D. Yilmazer, A Privacy-Aware Semantic Model for Provenance Management, in: Commun. Comput. Inf. Sci., 2014: pp. 162–169. doi:10.1007/978-3-319-13674-5_16.

[12] Dao, T.T., T.N. Hoang, X.H. Ta, and M.C. Ho Ba Tho, Knowledge-based personalized search engine for the Web-based Human Musculoskeletal System Resources (HMSR) in biomechanics, *J. Biomed. Inform.* **46** (2013) 160–173. doi:10.1016/j.jbi.2012.11.001.

[13] Dersingh, A., R. Liscano, and A. Jost, Context-aware Access Control Using Semantic Policies, *Spec. Issue Auton. Comput. Syst. Appl. ,Ubiquitous Comput. Commun. Journal, UBICC.* **ACSA-Spe** (2008) 19–32.

[14] Dixit, S., K.P. Joshi, and S. Geol Choi, Multi Authority Access Control in a Cloud EHR System with MA-ABE, *Proc. - 2019 IEEE Int. Conf. Edge Comput. EDGE 2019 - Part 2019 IEEE World Congr. Serv.* (2019) 107–109. doi:10.1109/EDGE.2019.00032.

[15] Dong, X., R. Samavi, and T. Topaloglou, COC: An ontology for capturing semantics of circle of care, *Procedia Comput. Sci.* **63** (2015) 589–594. doi:10.1016/j.procs.2015.08.389.

[16] Dridi, A., S. Sassi, and S. Faiz, Towards a semantic medical internet of things, *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA.* **2017-Octob** (2018) 1421–1428. doi:10.1109/AICCSA.2017.194.

[17] Durmuş, V., and M. Uydaci, A Legal Framework for Healthcare: Personal Data Protection for Health Law in Turkey, in: Handb. Res. Intrusion Detect. Syst., IGI Global, 2020: pp. 219–236. doi:10.4018/978-1-7998-2242-4.ch011.

[18] EUROPEAN PARLIAMENT, T., and T. COUNCIL OF THE EUROPEAN UNION, EU General Data Protection Regulation, 2016.

[19] Garde, S., P. Knaup, E.J.S. Hovenga, and S. Heard, Towards semantic interoperability for electronic health records: Domain knowledge governance for openEHR archetypes, *Methods Inf. Med.* **46** (2007) 332–343. doi:10.1160/ME5001.

[20] Grodzinsky, F.S., and H.T. Tavani, P2P networks and the Verizon v. RIAA case: Implications for personal privacy and intellectual property, *Ethics Inf. Technol.* **7** (2005) 243–250. doi:10.1007/s10676-006-0012-4.

[21] Häyrinen, K., K. Saranto, and P. Nykänen, Definition, structure, content, use and impacts of electronic health records: A review of the research literature, *Int. J. Med. Inform.* **77** (2008) 291–304. doi:10.1016/j.ijmedinf.2007.09.001.

[22] Health Level Seven International, HL7 FHIR Specification, *HL7 FHIR Specif.* (2019). http://hl7.org/fhir/%0Ahttp://hl7.org/implement/standards/fhir/.

[23] Health Level Seven International, HL7 FHIR Security, *HL7 FHIR Secur.* (2019). https://www.hl7.org/fhir/security.html (accessed April 19, 2021).

[24] HIMSS, H.I. and M.S.S., Definition of Interoperability, (2013) 2013.

[25] Huertas Celdrán, A., M. Gil Pérez, F.J. García Clemente, and G. Martínez Pérez, Preserving patients' privacy in health scenarios through a multicontext-aware system, *Ann. Des Telecommun. Telecommun.* **72** (2017) 577–587. doi:10.1007/s12243-017-0582-7.

[26] IHE, IHE ITI TF-1 (rev 17), *Int. J. Healthc. Technol. Manag.* **1** (2020) 1–177.

[27] Iroju, O., A. Soriyan, I. Gambo, and J. Olaleke, Interoperability in Healthcare: Benefits, Challenges and Resolutions, *Int. J. Innov. Appl. Stud.* **3** (2013) 262–270. http://www.ijias.issr-journals.org/abstract.php?article=IJIAS-13-090-01.

[28] Ismael Taher, K., R. Hasan Saeed, R. Kh. Ibrahim, Z. Najat Rashid, L. M. Haji, N. Omar, and H. Ismat Dino, Efficiency of Semantic Web Implementation on Cloud Computing: A Review, *Qubahan Acad. J.* **1** (2021) 1–9. doi:10.48161/qaj.v1n3a72.

[29] Joshi, K.P., Y. Yesha, and T. Finin, An Ontology for a HIPAA compliant cloud services, (n.d.).

[30] Karthik, N., and V.S. Ananthanarayana, An Ontology Based Trust Framework for Sensor-Driven Pervasive Environment, *AMS 2017 - Asia Model. Symp. 2017 11th Int. Conf. Math. Model. Comput. Simul.* (2018) 147–152. doi:10.1109/AMS.2017.31.

[31] Kirrane, S., S. Villata, and M. D'Aquin, Privacy, security and policies: A review of problems and solutions with semantic web technologies, *Semant. Web*. **9** (2018) 153–161. doi:10.3233/SW-180289.

[32] Kitsos, P., and P. Pappa, Mobile Communications Privacy, in: Encycl. Inf. Sci. Technol. Third Ed., IGI Global, 2015. doi:10.4018/978-1-4666-5888-2.ch202.

[33] Kondylakis, H., G. Flouris, I. Fundulaki, V. Papakonstantinou, and M. Tsiknakis, Flexible access to patient data through e-consent, *MOBIHEALTH 2015 - 5th EAI Int. Conf. Wirel. Mob. Commun. Healthc. - Transform. Healthc. through Innov. Mob. Wirel. Technol.* (2015). doi:10.4108/eai.14-10-2015.2261673.

[34] Li, Z., C.H. Chu, and W. Yao, A semantic authorization model for pervasive healthcare, *J. Netw. Comput. Appl.* **38** (2014) 76–87. doi:10.1016/j.jnca.2013.06.006.

[35] Lima, V.C., D. Alves, F.C. Pellison, V.T. Yoshiura, N.Y. Crepaldi, and R.P.C.L. Rijo, Establishment of access levels for health sensitive data exchange through semantic web, *Procedia Comput. Sci.* **138** (2018) 191–196. doi:10.1016/j.procs.2018.10.027.

[36] Lima, V.C., F.C. Pellison, F.A. Bernardi, D. Alves, and R.P.C.L. Rijo, Security Framework for Tuberculosis Health Data Interoperability Through the Semantic Web, *Int. J. Web Portals*. **13** (2021) 36–57. doi:10.4018/ijwp.2021070103.

[37] Liu, Z., and J. Wang, A fine-grained context-aware access control model for health care and life science linked data, *Multimed. Tools Appl.* **75** (2016) 14263–14280. doi:10.1007/s11042-016-3269-6.

[38] Lu, Y., and R.O. Sinnott, Semantic privacy-preserving framework for electronic health record linkage, *Telemat. Informatics*. **35** (2018) 737–752. doi:10.1016/j.tele.2017.06.007.

[39] Mamun, A. Al, K. Salah, S. Al-maadeed, and T.R. Sheltami, BigCrypt for Big Data Encryption, (2017) 93–99.

[40] Moher, D., A. Liberati, J. Tetzlaff, and D.G. Altman, Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine), *Ann. Intern. Med.* **151** (2009) 264–269. doi:10.1371/journal.pmed.1000097.

[41] Munn, Z., M.D.J. Peters, C. Stern, C. Tufanaru, A. McArthur, and E. Aromataris, Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach, *BMC Med. Res. Methodol.* **18** (2018) 1–7. doi:10.1186/s12874-018-0611-x.

[42] Nass, S.J., L.A. Levit, and O.L. Gostin, Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research., Washington, D.C., 2009. http://www.ncbi.nlm.nih.gov/books/NBK9571/ accessed 01/08/2016.

[43] National Institute of Standards and Technolgy, NIST Privacy Framework - a tool for improving privacy through enterprise risk management, 2020. doi:10.6028/NIST.CSWP.01162020.

[44] National Institute of Standards and Technology, Minimum Security Requirements for Federal Information and Information Systems, 2006. http://csrc.nist.gov/publications/fips/fips200/FIPS-200-final-march.pdf.

[45] Noor, S., M. Ahmed, M.N. Saqib, M. Abdullah-Al-Wadud, M.S. Islam, and Fazal-E-Amin, Ontology for attack detection: Semantic-based approach for genomic data security, *J. Med. Imaging Heal. Informatics*. **7** (2017) 1309–1323. doi:10.1166/jmihi.2017.2221.

[46] Papakonstantinou, V., G. Flouris, I. Fundulaki, and H. Kondylakis, Securing access to sensitive RDF data, *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. **8798** (2014) 455–460. doi:10.1007/978-3-319-11955-7_66.

[47] Peters, M.D.J., C.M. Godfrey, H. Khalil, P. McInerney, D. Parker, and C.B. Soares, Guidance for conducting systematic scoping reviews, *Int. J. Evid. Based. Healthc.* **13** (2015) 141–146. doi:10.1097/XEB.0000000000000050.

[48] Poulymenopoulou, M., F. Malamateniou, and G. Vassilacopoulos, An Access Control Framework for Pervasive Mobile Healthcare Systems Utilizing Cloud Services, *Int. ICST Conf. Wirel. Mob. Commun. Healthc.* **83** (2012) 380–385. doi:10.1007/978-3-642-29734-2_52.

[49] Rahmouni, H.B., M.C. Mont, K. Munir, and T. Solomonides, A SWRL bridge to XACML for clouds privacy compliant policies, *CLOSER 2014 - Proc. 4th Int. Conf. Cloud Comput. Serv. Sci.* (2014) 27–37. doi:10.5220/0004853900270037.

[50] Rahmouni, H.B., T. Solomonides, M.C. Mont, and Simon Shiu, Privacy compliance and enforcement on European healthgrids: An approach through ontology, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **368** (2010) 4057–4072. doi:10.1098/rsta.2010.0169.

[51] Raman, P., H.G.H. Kayacık, and A. Somayaji, Understanding Data Leak Prevention, *Annu. Symp. Inf. Assur.* **2016** (2011) 27–31. doi:10.1109/IConAC.2015.7313979.

[52] Rhayem, A., M.B.A. Mhiri, and F. Gargouri, Semantic Web Technologies for the Internet of Things: Systematic Literature Review, *Internet of Things (Netherlands)*. **11** (2020). doi:10.1016/j.iot.2020.100206.

[53] Robu, I., V. Robu, and B. Thirion, An introduction to the Semantic Web for health sciences librarians., *J. Med. Libr.*

Assoc. **94** (2006) 198–205.

[54] Stouffer, K., V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, Guide to Industrial Control Systems (ICS) Security, 2015. doi:10.6028/NIST.SP.800-82r2.

[55] Sun, L., J. Yong, and J. Soar, Access control management for e-Healthcare in cloud environment, *ICST Trans. Scalable Inf. Syst.* **1** (2014) e3. doi:10.4108/sis.1.2.e3.

[56] The Apache Software Foundation, Apache Jena - A free and open source Java framework for building Semantic Web and Linked Data applications, (2011). https://jena.apache.org/ (accessed April 8, 2020).

[57] Tiwari, S.M., S. Jain, A. Abraham, and S. Shandilya, Secure semantic smart healthcare (S3HC), *J. Web Eng.* **17** (2019) 617–646. doi:10.13052/jwe1540-9589.1782.

[58] Yarmand, M.H., K. Sartipi, and D.G. Down, Behavior-based access control for distributed healthcare environment, *Proc. - IEEE Symp. Comput. Med. Syst.* (2008) 126–131. doi:10.1109/CBMS.2008.14.

[59] Zenuni, X., B. Raufi, F. Ismaili, and J. Ajdari, State of the Art of Semantic Web for Healthcare, *Procedia - Soc. Behav. Sci.* **195** (2015) 1990–1998. doi:10.1016/j.sbspro.2015.06.213.

[60] Zissis, D., and D. Lekkas, Addressing cloud computing security issues, *Futur. Gener. Comput. Syst.* **28** (2012) 583–592. doi:10.1016/j.future.2010.12.006.

Table 2. Details of the selected studies

		Features								
Ref.	**Article Title**	**Confidentiality**	**Integrity**	**Availability**	**Privacy**	**Provenance**	**Authentication**	**Authorization**	**Category**	**Underlying mechanisms/technologies**
[38]	A fine-grained context-aware access control model for health care and life science linked data	-	-	-	-	-	x	x	Access Control	- XACML - OWL Ontology - SWRL rules - Inference/reasoning engine
[34]	A semantic authorization model for pervasive healthcare	-	-	-	-	-	-	x	Access Control	- Ontology - Inference/reasoning engine
[50]	A SWRL bridge to XACML for clouds privacy compliant policies	-	-	-	x	-	x	x	Access Control	- XACML - OWL Ontology - SWRL rules
[56]	Access control management for e-Healthcare in cloud environment	-	-	-	-	-	x	x	Access Control	- XACML - OWL Ontology - RBAC security model
[49]	An access control framework for pervasive mobile healthcare systems utilizing cloud services	-	-	-	-	-	-	x	Access Control	- OWL Ontology - SWRL Rules - Inference/reasoning engine - RBAC security model
[59]	Behavior-Based Access Control for Distributed Healthcare Environment	-	-	-	x	-	-	x	Access Control	- Ontology - Inference/reasoning engine
[13]	Context-aware Access Control Using Semantic Policies	-	-	-	-	-	-	x	Access Control	- XACML - OWL Ontology - SWRL rules - Inference/reasoning engine
[35]	Establishment of access levels for health sensitive data exchange through semantic web	-	-	-	-	-	x	x	Access Control	- Ontology - Jena - JSON-LD - Semantic API
[33]	Flexible access to patient data through e-Consent	-	-	-	x	x	-	x	Access Control	- Ontology - Abstract annotation model - RDF
[14]	Multi Authority Access Control in a Cloud EHR System with MA-ABE	x	-	-	x	-	x	x	Access Control	- Encryption (MA-ABE) - ABAC security model - Ontology - SWRL rules - Inference/reasoning engine
[51]	Privacy compliance and enforcement on European healthgrids: An approach through ontology	-	-	-	x	-	-	x	Access Control	- OWL Ontology - SWRL rules
[47]	Securing access to sensitive RDF data	-	-	-	x	x	-	x	Access Control	- Ontology - Abstract annotation model - RDF
[39]	Semantic privacy-preserving framework for electronic health record linkage	-	-	-	x	-	-	x	Access control	- XACML - OWL Ontology - SWRL rules - Inference/reasoning engine

Table 2. Details of the selected studies

		Features								
Ref.	**Article Title**	**Confidentiality**	**Integrity**	**Availability**	**Privacy**	**Provenance**	**Authentication**	**Authorization**	**Category**	**Underlying mechanisms/technologies**
[16]	Towards a semantic medical internet of things	x	-	-	x	-	x	x	Access Control	- Ontology - RDF - IoT devices
[5]	Using OWL and SWRL to represent and reason with situation-based access control policies	-	-	-	-	-	-	x	Access Control	- OWL Ontology - SWRL rules - Description Logic - Inference/reasoning engine
[8]	ARTEMIS: towards a secure interoperability infrastructure for healthcare information systems	x	x	-	x	-	x	x	Interoperability Infrastructure	- Semantic Web Services - Ontology - Inference/reasoning engine - Encryption (Triple-DES)
[58]	Secure semantic smart healthcare (S3HC)	x	x	x	x	-	x	x	Interoperability Infrastructure	- Ontology - SWRL rules - Inference/reasoning engine - SPARQL - RDF/XML Security - Encryption - Hash functions - IoT devices
[36]	Security Framework for Tuberculosis Health Data Interoperability Through the Semantic Web	x	x	x	-	-	x	x	Interoperability Infrastructure	- Ontology - Hybrid cryptography - Hash functions - Jena - SPARQL - D2Rq Server - HyperGraphQL - Semantic APIs - JSON-LD - RDF
[9]	A Model-driven Privacy Compliance Decision Support for Medical Data Sharing in Europe	-	-	-	x	-	-	-	Privacy Compliance	- Ontology - SWRL rules - Inference/reasoning engine
[2]	An integrated framework for privacy protection in IoT — Applied to smart healthcare	-	-	-	x	-	-	-	Privacy Compliance	- OWL Ontology - Inference/reasoning engine - IoT devices
[29]	An Ontology for a HIPAA compliant cloud services	-	-	-	x	-	-	-	Privacy Compliance	- OWL Ontology
[15]	COC: An ontology for capturing semantics of circle of care	-	-	-	x	x	-	x	Privacy Compliance	- Ontology - RDF - SPARQL
[10]	Improving privacy in health care with an ontology-based provenance management system	-	-	-	x	x	-	-	Privacy Compliance	- Ontology
[12]	Knowledge-based personalized search engine for the Web-based Human Musculoskeletal System Resources (HMSR) in biomechanics	x	-	-	-	-	-	-	Privacy Compliance	- OWL Ontology - Multi-agent semantic crawler for the Web - Asymmetric encryption

Table 2. Details of the selected studies

		Features								
Ref.	**Article Title**	**Confidentiality**	**Integrity**	**Availability**	**Privacy**	**Provenance**	**Authentication**	**Authorization**	**Category**	**Underlying mechanisms/technologies**
[45]	Ontology for Attack Detection: Semantic-Based Approach for Genomic Data Security	-	-	-	x	-	-	-	Privacy Compliance	- OWL Ontology - SWRL rules - Jena - Inference/reasoning engine
[25]	Preserving patients' privacy in health scenarios through a multicontext-aware system	-	-	-	x	-	-	-	Privacy Compliance	- OWL Ontology - SWRL rules - Jena - Inference/reasoning engine