

Received 20 August 2022, accepted 5 September 2022, date of publication 8 September 2022, date of current version 15 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3204970

RESEARCH ARTICLE

DICON: A Domain-Independent Consent Management for Personal Data Protection

EMRE OLCA¹ AND OZGU CAN²

¹Department of Software Engineering, Maltepe University, 34857 Istanbul, Turkey

²Department of Computer Engineering, Ege University, 35100 Izmir, Turkey

Corresponding author: Ozgu Can (ozgu.can@ege.edu.tr)

ABSTRACT The development of technology accelerated the digital transformation of information systems. As a consequence of this digitization, data became available at any time and in any place. However, despite this ease of data accessibility, persons' privacy concerns and threats to data privacy have emerged. Thus, serious privacy problems arise while collecting, storing, accessing, sharing, and archiving personal data. Consent management aims to prevent these problems by preserving privacy and protecting personal data. Hence, there are international treaties and legal regulations for personal data protection which state that consent is required to collect, store, manage and share personal data. In this study, a Semantic Web-based personal consent management model is proposed to protect personal data privacy. The proposed model is domain-independent and aims to control and manage the consent of a person. In order to provide the privacy protection of personal data, the proposed model allows individuals to establish their privacy preferences by determining who can access their personal information, for what purposes, and under what circumstances. For this purpose, a group of ontology is created to ensure the informed consent process. The proposed consent management model is generic. As similar to general personal information, personal health information is also sensitive and must be protected from data leakage. Therefore, the proposed generic model is implemented with Semantic Web technologies and demonstrated for the healthcare domain.

INDEX TERMS Consent, data protection, knowledge-based systems, knowledge representation, ontology, privacy, Semantic Web.

I. INTRODUCTION

The broad usage of the Internet has increased the volume of personal data that are stored and processed by various information systems. Personal data can be used to identify an individual directly from a data source or by linkage of information which gathers data from different data sources and links data across data sources. The General Data Protection Regulation (GDPR) which is the regulation in European Union (EU) on data protection and privacy defines personal data as “personal data means any information relating to an identified or identifiable natural person (data subject)” [1]. The regulation also states that “personal data which are, by their nature, particularly sensitive in relation to fundamental rights and freedoms merit specific protection as the

context of their processing could create significant risks to the fundamental rights and freedoms”. Hence, this sensitive data must be protected from unauthorized access to protect personal privacy. Therefore, the data subject must be aware of who controls her data and for what purposes her data will be processed. For this purpose, consent is obtained from the data owner to use her personal data for different purposes. Consent means the decision of the user about how her personal data is accessed, managed, and shared. GDPR defines consent as “any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her” [1]. Thus, consent establishes a legal basis for protecting the data owner's personal data. Therefore, consent management becomes a necessary challenge in privacy preservation. Consent management controls the access management process

The associate editor coordinating the review of this manuscript and approving it for publication was Sedat Akleylek.

of personal data by considering consents that are granted by users. In [2], consent management is stated as one of the most important concepts in the Future Internet. Furthermore, it is stated that the Future Internet is going to see a lot of lawsuits that deal with what service a person has consented to, what scope that consent had, and how the service has evolved after the consent was given. Therefore, obtaining the consent of the user and managing the process of this consent is very important for the security and privacy of personal data.

The consent information needs to be maintained and shared by multiple parties such as the data subject, the data controller, data processor, and data protection authorities [3]. Hence, the consent information needs to be represented in a meaningful way to provide interoperability between these parties. Also, users must be able to create, revoke and update their consents. Consequently, an efficient consent management system has inevitable importance to preserve user privacy. For this purpose, a domain-independent consent management model for preserving personal data privacy is proposed in this study. The proposed consent management model is based on the Semantic Web which is an extension of the current web. In the Semantic Web, information is given in a well-defined meaning to represent information more meaningfully for both humans and computers [4]. Thus, the Semantic Web provides interoperability which is the ability to exchange information and to use this information between systems. In the Semantic Web, ontologies are used to provide the semantic description of information. Therefore, ontologies are considered as the backbone of the Semantic Web [5]. An ontology is defined as an explicit specification of a conceptualization [6]. Ontologies are used to represent the domain knowledge by describing the concepts and relationships between these concepts within the related domain. In this study, ontologies are used to represent a generic consent management system. The proposed ontology-based personal consent management model allows individuals to establish their privacy preferences to determine who can access their personal information, for what purposes, and under what circumstances. Therefore, an ontology-based consent management approach allows to represent the high-level knowledge of the consent management mechanism and to define the formal representation of the consent management model.

In terms of data protection regulations, consent management has a significant role in protecting users' privacy. United Nations Conference on Trade and Development (UNCTAD) states that *128 out of 194 countries had put in place legislation to secure the protection of data and privacy* [7]. In this context, a Law on the Protection of Personal Data (LPPD) is also enacted in Turkey [8]. The purpose of this law is to protect the fundamental rights and freedoms of people, particularly the right to privacy, with respect to the processing of personal data. LPPD is based on EU's Data Protection Directive 95/46/EC which was repealed when GDPR is entered into force. The differences between LPPD and GDPR are presented in [9]. The proposed personal

consent management model is based on LPPD. Hence, Personal Consent Ontology (PCO) and Domain-Independent Consent Management Ontology (DICON) are created for the proposed personal consent management model. These ontologies are based on LPPD [8] and to the best of our knowledge there is no other ontology that is based on this law. Also, FOAF (Friend-Of-A-Friend) [10] and Relationship Ontology [11] are integrated with the proposed personal consent management model. FOAF and Relationship Ontology are existing ontologies in the literature. FOAF defines a FOAF profile to link people by defining relationships between people. Relationship Ontology defines a vocabulary to describe relationships between people. In this study, both of these ontologies are extended according to the needs of the personal consent management model and imported into the DICON.

Privacy is defined as the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others [12]. Privacy protection ensures that data is not obtained, used, or disclosed by unauthorized persons. For this purpose, individuals must have control over their personal information to determine who can access it. In order to protect privacy, it should be ensured that personal information is not shared with third parties without the explicit consent of the data owner. With the dynamics of today's new technological developments and disruptive business models, a huge amount of personal data is being collected, used, analyzed, and exchanged. As a result of these developments, an increasing number of accidental or intentional data breaches occur. Thus, the demand for data privacy has grown due to the right to control how personal information is collected, with whom it is shared, and how it is used. Therefore, the demand for data privacy is provided with legal regulations. Privacy regulations provide the legal bases for data processing by defining consent in the context of data privacy and data rights. Consent is the processing of personal data. Hence, consent is the key to ensuring privacy for individuals' data and addressing individuals' data privacy requirements. In this manner, consent management is an important mechanism to ensure enhanced privacy.

The goal of this study is also to create a domain-independent personal consent management model for the privacy protection of personal data. Therefore, after creating domain-independent ontologies a case study is conducted to validate the proposed model. As consent management has an important role in the healthcare domain due to the raising privacy concerns of patients, the use case of the presented domain-independent consent management model is performed on the healthcare domain. In healthcare, consent management enables patients to specify their consents to allow or to deny third parties access, to process, and to share their health information. Thus, patient privacy is ensured in accordance with the patient's consent policy. In order to perform the case study, the e-Pulse system (also known as e-Nabiz) which is a trusted personal health record system of

Turkey [13] is chosen as the domain of the proposed consent management model. The e-Pulse system allows citizens and health professionals to access and manage health data that is collected from health institutions. For the use case study, e-Pulse Ontology is created and used as the domain ontology. Also, e-Pulse Ontology is the first ontology that is based on the e-Pulse personal health record system. Furthermore, PurposeOfUse Ontology [14] of HL7 that represents the access purposes for the healthcare domain is integrated with the e-Pulse Ontology.

The main contributions of this study are as follows: (i) the first ontology-based consent management model that is based on Personal Data Protection Law [8] is introduced; (ii) a personalized privacy approach is presented to preserve privacy and to support personalized privacy; (iii) a Domain-Independent Consent Management Model (DICON) is introduced and a Domain-Independent Consent Management Ontology is created; (iv) a well-defined ontological model of the consent management process is provided and a Personal Consent Ontology (PCO) is created; (v) e-Pulse Ontology that is based on the e-Pulse personal health record system [13] is created; (vi) FOAF [10] and Relationship Ontology [11] are extended and integrated with the proposed consent management model.

The structure of the paper is organized as follows. Section 2 presents the current status of the literature. Section 3 clarifies the consent management process for data protection. Section 4 explains the proposed personalized consent management model. Section 5 describes the ontologies that are created for the proposed model. Section 6 presents a case study to illustrate the proposed domain-independent model for the healthcare domain. Finally, Section 7 concludes and describes the future work.

II. RELATED WORK

The consent principle is widely covered by personal data protection laws to protect the privacy of sensitive information according to the privacy preferences of the owner of the sensitive information. The term “informed consent” is one of the characteristics of consent and is used as a requirement for the consent [15]. In informed consent, consent must be meaningful and the individual needs to understand its implications [15]. Informed consent was first filed in a court case in California in 1957 [16]. Afterward, it became a compulsory practice in many fields, especially in the healthcare domain. The American Medical Association emphasizes the need to obtain informed consent in medical treatment and indicates that informed consent is a fundamental process for both ethics and law [17].

Consent management is a process that manages access to personal data according to the permissions granted by the individual. Electronic consent management holds the person's consent, manages the process, permits, or denies the access requests to the personal data according to the consent that is given by the person. The process of an e-consent is presented in [15]. In [18], a patent is filed for a system

and method to manage consent requests for an enterprise and to provide consent-controlled data access among enterprises. As the importance of privacy and data protection is increasingly recognized, countries have adopted data protection regulations on the processing of personally identifiable data such as GDPR [1], HIPAA [19], GLBA [20] or LPPD [8]. Thereupon, the consent management process allows systems to meet these regulations by obtaining user consent and enables systems to be compliant with regulations. Hence, in various studies, these regulations are used as a legal basis for the management of the consent. In [21], an agile tool for consent control is proposed to facilitate compliance with the EU's GDPR. The proposed tool provides interaction between data subjects, controllers, and processors. In [22], a formal framework for consent management in line with GDPR is designed and a formalization of runtime policy compliance is presented.

The Semantic Web enables to build the model of a domain and represents information in a more meaningful and machine-understandable format. Thus, the Semantic Web enables computers and people to work in cooperation [4]. For this purpose, ontologies are used to enhance the functioning of the Semantic Web. An ontology defines terms, their definitions, axioms, and relationships between terms. Therefore, ontology is used to describe various aspects of the domain being modeled [23] and to provide an explicit specification of the domain information to be reused across various applications. In the literature, consent management is one of the application areas of Semantic Web technologies. In [24] and [25], semantic models for personal consent management are proposed. These models are generic representations of the consent management model and are not based on any legal basis. A privacy ontology that models the main concepts of GDPR is presented in [26].

Privacy-preserving consent management for the Internet of Things is proposed in [27]. The proposed solution is based on cryptographic consent proof issued by users. The key principles of consent and the challenges raised by pervasive systems are discussed and a set of recommendations to designers are presented in [28]. As the consent information needs to be managed in various information systems, most of the consent management studies in the literature are based on healthcare information systems. The effective coordination of healthcare relies on the communication of patients' confidential information between different health and community care services [29]. Thus, it must be ensured that the patient's privacy is preserved according to the consent policy of the patient. For this purpose, design principles of an e-consent system in healthcare are presented in [29] and a consent-based authorization architecture for health services is proposed in [30]. In [31], an informed consent-based access model is proposed to share drug information that is stored in electronic health records in Norway. A Regional Health Information Network is presented in [32]. The presented regional network is in compliance with German legislation and relies on the patient's consent to share documents and medical

data with other care delivery organizations. The current and future state of consent management in healthcare is discussed in [33]. In [34], digital rights management is combined with electronic consent and workflow-based access control for a secure healthcare information system. In [35], an ontology-based model of subject's permissions is proposed to define permissions that are obtained from individuals who signed the informed consent form. Also, ontology-based reasoning is provided and the presented model is evaluated for collecting and storing biospecimens for use in cancer research. An ontology-based framework is developed and a prototype that is using the open-source Electronic Medical Record system is presented in [36]. In [37], an Informed Consent Ontology (ICO) is developed to support the data integration and inference in the clinical study research domain. The Vaccination Informed Consent Ontology (VICO) presented in [38] is extended from ICO and focuses on vaccination screening questionnaire in the vaccination informed consent domain.

Personal Consent Ontology is an ontology that is created by analyzing LPPD. In the literature, there are different ontologies to ensure the privacy and security of personal data. Among these ontologies, the closest ontological models to our approach are HL7 Security and Privacy Ontology [39] and Ontology for the EU's General Data Protection Regulation [40].

Security and Privacy Ontology is developed as a Health Level 7 Standard [41], [42] to improve international health standards, focus on the security and privacy aspects of the healthcare domain, and manage individuals' permissions. The ontology aims to support descriptions of security policies, privacy policies, and consent directives [39]. HL7 standards focus on information exchange, association, sharing, and usage among independent computer systems such as hospital information systems, laboratory systems, pharmacy systems, and corporate systems. In this context, Security and Privacy Ontology is also developed especially for the healthcare domain. The ontology includes terms related to healthcare and defines relationships between concepts under the healthcare domain. The PCO and Security and Privacy Ontology have common classes, such as Person, Object, Purpose of Use, Sensitivity, and Organization. Whereas, the proposed PCO is a domain-independent and generic consent ontology. Thus, the PCO is a high-level ontology that integrates all domain ontologies and it can be used in any domain where user consent is obtained and consent management is needed. Therefore, the PCO can be used in various domains. As consent is necessary information that exists in several domains, the PCO can be used not only in the healthcare domain but also in every domain where personal data is stored.

The Ontology for the EU's General Data Protection Regulation is developed for GDPR. The ontology contains classes that define GDPR-related regulations and security controls. However, the ontology does not include classes and relations that can operate based on the consent

management process. Also, key concepts of consent management such as Personal Data, Person, Consent, Data Processor, Transaction, and Purpose are not defined in the ontology. The proposed PCO includes the related concepts to support the consent management process. As the Ontology for GDPR defines legal system elements, the PCO defines concepts and relationships to perform the consent management process. As seen, the PCO is more comprehensive than the existing ontologies. Moreover, the proposed ontological model is domain-independent and is not developed for a single domain. Thus, it can be used in any domain where consent is needed.

Different from the existing works, the proposed DICON model integrates the concept of consent with the privacy requirements to fulfill individuals' privacy needs by considering the legal bases and interoperability between systems. For this purpose, DICON is developed based on the LPPD [8] and Semantic Web technologies. Therefore, a group of consent management ontology is presented to ensure the privacy of personal data within the scope of the proposed DICON model. Also, personal information and the relationship between individuals are represented with FOAF profiles [10] and Relationship Ontology [11] to provide a fully semantic consent management framework. Further, an ontology is developed for the e-Pulse personal health record system of Turkey [13] that stores patients' health information to evaluate the proposed model for the healthcare domain. Moreover, the PurposeOfUse Ontology [14] that is created by HL7 is used to represent the purpose of accessing data in the healthcare domain. Therefore, the presented model ensures that both e-pulse data and consent data are consistent, well-formulated, semantically rich, and traceable. Thus, the proposed approach provides a well-defined model of the consent management process. Also, the proposed personalized consent management model is domain-independent. Hence, the proposed model can be used in various domains.

III. CONSENT MANAGEMENT FOR PERSONAL DATA PROTECTION

LPPD provides a holistic approach for securing the privacy of personal data. According to the Law, the consent of the person is essential for the processing of personal data. In the literature, consent is a general term under data privacy. Data privacy is concerned with the handling, processing, usage, and storage of personal data. Data privacy is defined by LPPD as the inability to obtain, use or disclose data by unauthorized persons. Thus, data privacy cannot be guaranteed unless personal data is protected. Consent is used to ensure the privacy of personal data. OECD guidelines on data protection state that personal data should be limited to the collecting and processing for the confidentiality of personal data [43]. According to the European Convention on Human Rights [44], the European Union Convention on Fundamental Rights [45] and the Council of Europe [46], the consent of the person is required to ensure the confidentiality of personal data. As seen, the consent principle is widely covered by

several privacy legislation, laws, and treaties to limit the discovery of sensitive information according to the data owner's permission. LPPD uses the term "Open Consent" instead of the term "Consent". Open Consent states that consent on a particular subject is based on the information and free consent; and also the processing of personal data can only take place with the consent of the person. Consent management handles the access control request, and permits or prohibits the access request based on the person's open consent.

LPPD describes actors of the consent management and relationships between them. In this study, all possible elements of the consent management system are extracted from LPPD. As a result, actors, objects, and processes within the proposed consent management system are as follows: Personal Data, Relevant Person, Open Consent, Processing of Personal Data (*obtaining, saving, storing, maintaining, changing, rearranging, explaining, transferring, taking over, being available, prevention, classification, or prevention of use*), President, Personal Data Protection Board, Personal Data Protection Agency, Data Record System, Data Processor, Processing Conditions, Data Responsible, Real Person, Legal Entity, Special Personal Data, Data Transfer, Data Responsibilities Registry, Anonymization, Personal Data Deletion, Destruction.

The relationship between Data Controller, Data Processor and Data Registry System is shown in Figure 1. LPPD defines the main members and relations between these members as follows [8]:

- Personal Data is any information that is related to an identifiable person.
- Explicit Consent is an informed and a free-willed consent on a specific subject.
- Data Subject is the real person whose Personal Data is processed; the owner of the Personal Data.
- Data Controller is a natural or legal person who is responsible for setting up and managing the data logging system, determining the means of the processing of the personal data. Data Controller is responsible for establishing and managing the Data Registry System and can anonymize, destroy and/or delete the Personal Data that is held by the Data Registry System. Data Controller may also transfer Personal Data to Foreign/Third Parties on the basis of the processing of the Personal Data. Data Controller authorizes Data Processor to process the Personal Data.
- Data Processor is a natural or legal person who processes the Personal Data on her behalf based on the authority conferred by the data officer. The authorized Data Processor processes the Personal Data by adhering to the Processing Conditions.

- Data Registry System is the recording system in which personal data is configured and processed according to certain criteria. It keeps the Personal Data. The Explicit Consent of the person should be checked to process the Personal Data. Personal Data may be also transferred to third parties or to abroad. In order to transfer the Personal Data to the third parties or to abroad, the Explicit Consent of the person or Other Cases that are specified in the law or the permission of the board is required.
- Processing Conditions are conditions of the processing of the Personal Data. In the presence of certain conditions, the Personal Data is processed without seeking the explicit consent of the person.
- Other Cases indicate that the processing of the Personal Data is not possible without expressing the consent of the person.

The workflow of the consent management process is shown in Figure 2. As seen in the figure, there are three processes in this workflow. Process 1 is the consent creation and the algorithm of this process is given in Algorithm 1. In Process 1, the Data Subject first creates a consent record for the Personal Data that is kept in the Data Registry System. While creating the consent record, Data Subject specifies for what purpose and action her data can be used. In addition, Data Subject selects the Sensitivity Level. Finally, Data Subject creates consent data by giving the consent information for the data, process, purpose, and sensitivity level that are chosen by the Data Subject. Thus, the scope of the Data Subject's selected criteria specifies whether the Personal Data is allowed to be accessed or not. In Process 2, the consent control is performed to process access to Personal Data. The data to be accessed, the purpose of use of the data, and the type of operation are selected and their consent information is checked. If there is no Explicit Consent, then Other Cases that are specified in LPPD are controlled. The process of transferring personal data is shown in Process 3. The transfer process is similar to Process 2. First, it is checked whether the person has given consent for the transfer of her data or not. If there is no consent given, then the Other Cases that are specified in LPPD is checked. If the required permission is not found, the permission of the board which controls the protection of personal data is checked based on the LPPD. The data transfer process starts if one permission from one of these three controls exists.

In the proposed DICON model, the relevant person can change her consent at any time. Changing or updating the consent information only affects the subsequent access. Thus, the related change/update does not affect the previous accesses. The withdrawal of the consent is controlled via Status class. The person can deactivate the consent record at any time. In addition, the partial acceptance process of the consent is also evaluated in reverse. Instead of partial acceptance, a personalized consent process is carried out on an atomic basis. Consent Data and Consent Policy

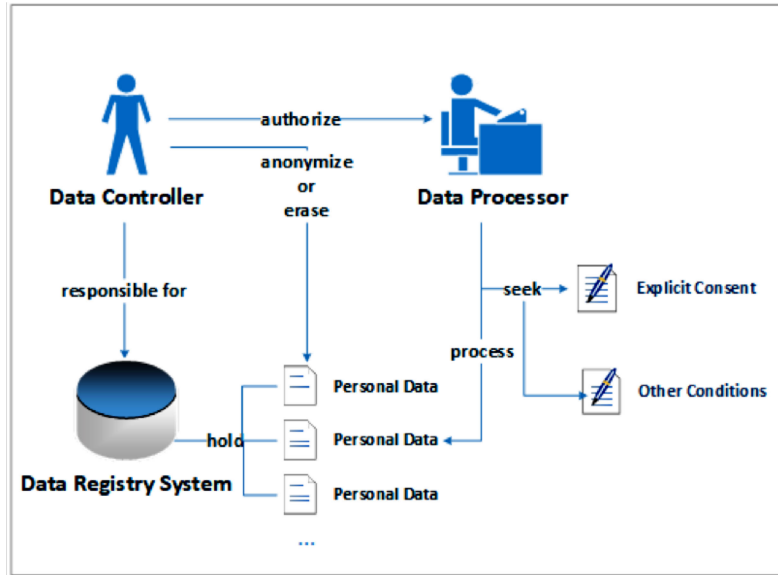


FIGURE 1. The relationship between the Data Controller, Data Processor and Data Registry System.

classes have a multi-class structure. Therefore, these classes define consent for the required amount of data with sensitivity levels. For example, while full consent can be given for the demographic data, a single consent can be given for a single data within the demographic data. Thus, consent can be defined for the requested data in a data set. This structure also supports partial consent from the perspective of giving consent. However, for partial acceptance of the consent requested by the data processor, an information request can be sent to the data owner via the application using the model, and the data owner can provide appropriate consent for the related request.

A. CONSENT MANAGEMENT ONTOLOGIES

The structure of the proposed consent management ontological model is presented based on Meta Object Facility (MOF) [48]. MOF is a metamodeling standard developed by the Object Management Group (OMG) for model-based engineering. MOF provides the basis for metamodel definition in OMG's family of Model Driven Architecture (MDA) languages [48]. As MOF is an extensible model-based integration framework for defining, processing, and combining metadata and data in a platform-independent manner [49], MOF can be used as a model to describe information patterns. In this context, the MOF model is called a meta-metamodel.

MOF defines a metadata architecture in four layers. The layers of the metamodel hierarchy are as follows: (i) M3 meta-metamodel layer, (ii) M2 metamodel layer (iii) M1 model layer, and (iv) M0 data layer. A model that represents a modeling language is called a metamodel [50]. Metamodels are used to identify models. Therefore, in the layered structure, there are metamodels that define a model in the upper layer of each model. The ontological hierarchy of the

Algorithm 1 CreateConsent

Input: Parameter $\alpha = (\alpha_1, \dots, \alpha_n)$ where α is the purpose for processing personal data D , parameter $\beta = (\beta_1, \dots, \beta_n)$ where β denotes operation on personal data D , Υ is the possible consent types, DS is data subject, CG is consent giver, PR is process, S is the sensitivity level which can be Low, Medium, High or Critical

Output: Consent C

```

1:  $C \leftarrow \text{Consent}()$ 
2:  $CD \leftarrow \text{ConsentData}()$ 
3: for each purpose  $P \in (\alpha_1, \dots, \alpha_n)$  do
4:    $P \leftarrow \alpha$ 
5: for each process  $PR \in (\beta_1, \dots, \beta_n)$  do
6:    $PR \leftarrow \beta$ 
7:  $S \leftarrow \text{SensitivityLevel}(\text{Low}, \text{Medium}, \text{High}, \text{Critical})$ 
8:  $CD = \text{ConsentData}(D, S, DS)$ 
9:  $\Upsilon \leftarrow \text{ConsentTypes}(\text{True}, \text{False})$ 
10: if (age < 18) then
11:    $CG \leftarrow \text{Parent}$ 
12: else
13:    $CG \leftarrow DS$ 
14: return  $C \leftarrow \text{Consent}(CG, CD, P, PR, \Upsilon)$ 
  
```

proposed study based on the MOF model is shown in Figure 3. Ontologies of the DICON are at M2, M1 and M0 layers of the MOF model.

As seen in Figure 3, the PCO is at the M2 layer and it is a super-model of the model at the M1 layer. The M2 layer also includes FOAF Ontology and Relationship Ontology. These three ontologies together provide super-classes and properties as meta-models. As the DICON integrates ontologies of the M2 layer into a single ontology, it is represented at the M1

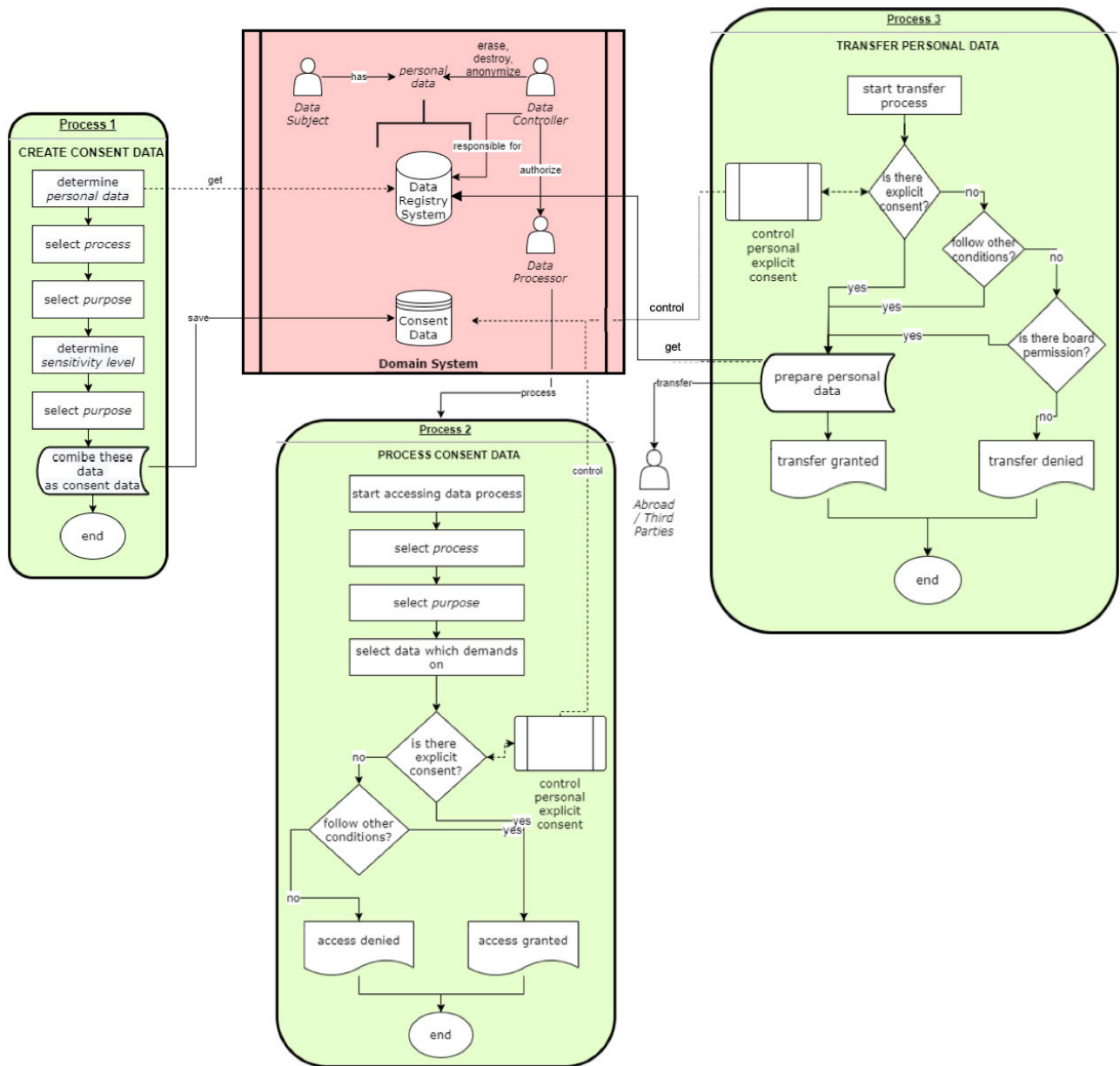


FIGURE 2. The flow diagram of the consent management process.

layer. The M0 layer is the runtime layer and individuals are represented at this level. Therefore, domain ontologies, E-Pulse Ontology and PurposeOfUse Ontology, are at the M0 layer.

The Protégé Ontology editor [51] is used to create the consent management ontologies of the DICON model. Protégé is an open-source ontology editor and framework to build knowledge-based solutions that was developed by the Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine.

B. PERSONAL CONSENT ONTOLOGY

Consent is a necessary information not only for a single domain, but also for various domains. Hence, in every

system where personal data is stored, the consent management process is mandatory in accordance with regulations such as LPPD and GDPR. Therefore, consent management should be applicable in every domain. For this purpose, the proposed model includes the Personal Consent Ontology (PCO). PCO is a meta-ontology that represents domain independent consent classes, the class hierarchy, properties and relationships between them.

The PCO is created by analyzing the LPPD text according to the Term Frequency-Inverse Document Frequency (TF-IDF) method [52]. In the literature, natural language processing, statistical analysis and machine learning techniques are used for the extraction of terminology from a text. For this purpose, statistical frequency methods such as

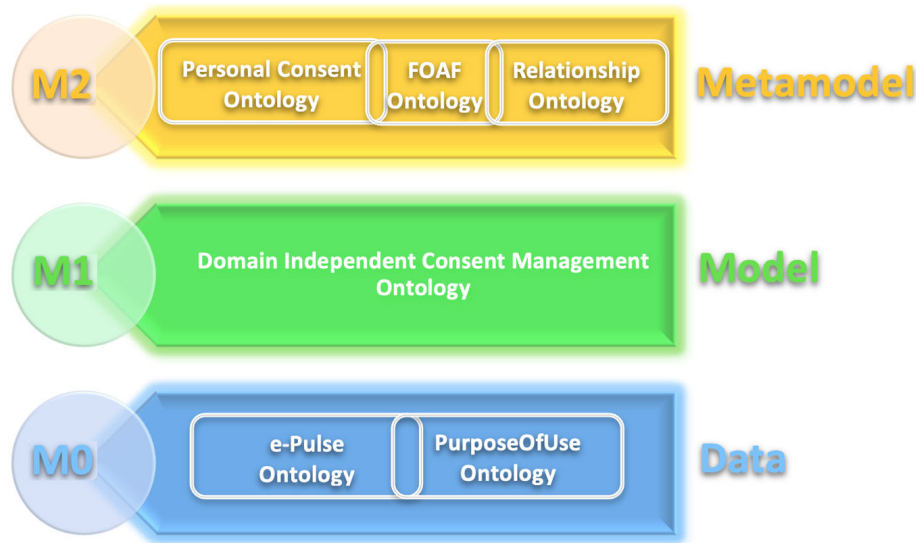


FIGURE 3. The ontology hierarchy of DICON.

TF-IDF [53], C-value [54], Named Entity Recognition (NER) [55], use of existing domain dictionaries or ontologies [56], “chunking” such as syntactic parsing [57] and design-based parsing such as the Hearst model [58] are frequently used. TF-IDF is a statistical calculation method that is used in natural language processing, text mining and information retrieval studies. TF-IDF method presents the term frequencies in a text by using two individual metrics, Term Frequency (TF) and Inverse Document Frequency (IDF), respectively. TF indicates the weight of a term in a document and IDF indicates the ratio of the total number of documents to the total number of documents containing the specified term and measures. Thus, TF measures how frequently a term occurs in a document and IDF measures the importance of the term. Therefore, TF-IDF method examines the key words in a document and measures the importance of terms within a document. Also, tools such as TerMine [59], Text2Onto [58], KIM Platform [60] are used to develop and/or support the ontology development process. In this study, the TF-IDF text mining method is used to generate the PCO from the LPPD. The formula that is used to determine the terms of the ontology is presented in the Equation (1):

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (1)$$

In the Equation (1), $tf_{i,j}$ is the number of occurrences of the term i in j , N is the total number of documents, and df_i is the number of different documents in which the related term (i) is occurred. The details of the term extraction are presented in [52].

According to the determined terms, classes defined in the PCO are as follows: The Personal Data is any information related to an identifiable real person and it is the subclass of the Object class; the Data Subject is the real person whose Personal Data is processed and it is

the subclass of the Person class; the Data Registry System is a registration system where the Personal Data is stored and processed according to the certain criteria; the Organization corresponds to any organization, institution and/or legal entity; the Data Processor is a natural or legal person who processes the Personal Data on her behalf on the basis of the authority granted by the Data Controller; the Data Controller is the natural or legal person who is responsible for establishing and managing the Data Registry System that determines the processing purposes of the Personal Data; the Authority is a class defined under the legal entity class which is the subclass of the Organization class and it is responsible for ensuring that Personal Data is processed in accordance with the fundamental rights and freedoms; the Purpose is an intention for the Object to be accessed; the Process is the action that will be performed on the Personal Data such as transfer, archive, store, update and save; the Sensitivity Level maintains the level of the sensitiveness that the Data Subject requires for her data and it has four sub-classes: Low, Medium, High and Critical; the Consent specifies the consent of the Data Subject for her Personal Data and it has two sub-classes: Personal Consent and Legal Consent; Personal Consent is the consent given by the Data Subject; the Legal Consent is the legal approval; Permission and Prohibition classes are the two common sub-classes of Personal Consent and Legal Consent; Status indicates whether the consent given is active or passive; Active and Passive classes are the sub-classes of Status; the Consent Policy is the policy class that corresponds to the directive which includes criteria such as the purpose, process, consent (permission or prohibition) that are determined by the Data Subject for her Personal Data. The Consent Policy is a

TABLE 1. The triple definitions of PCO.

| Subject | Predicate | Object |
|----------------|------------------------|--------------------|
| ConsentData | hasSensitivityLevel | SensitivityLevel |
| ConsentData | holdSubject | PersonalData |
| ConsentData | holdObject | DataSubject |
| ConsentData | holdDataRegistrySystem | DataRegistrySystem |
| ConsentPolicy | hasConsenter | DataSubject |
| ConsentPolicy | hasConsentData | ConsentData |
| ConsentPolicy | hasProcess | Process |
| ConsentPolicy | hasPurpose | Purpose |
| ConsentPolicy | status_type | Status |
| DataController | authorize | DataProcessor |
| DataController | hold | PersonalData |
| DataController | responsibleFor | DataRegistrySystem |
| DataSubject | hasObject | Object |

multi-element class and its elements are: Consent Data, Data Subject, Personal Consent, Purpose and Process. Multi-element classes reduce the complexity by keeping critical and interrelated classes together. Thereby, Consent Data is also a multi-element class that holds the Data Subject and the Personal Data together with the Sensitivity Level. The RDF triple definitions of the object properties that are defined in the PCO are given in Table 1. Figure 4 shows the visualization of the PCO.

C. DOMAIN INDEPENDENT CONSENT MANAGEMENT ONTOLOGY

The PCO is a meta-ontology that defines the meta-data required to define the Domain Independent Consent Management Ontology (DICMO). Thus, DICMO uses classes and properties that are defined in the PCO. The consent policy definitions and access decisions are represented in the DICMO. DICMO could also be customized according to the related domain. Therefore, new classes and properties could be added to the ontology according to the needs that would arise in the related domain. Moreover, DICMO imports FOAF and Relationship Ontology to store the personal information of the concerned person. It can also import different ontologies according to the related domain. While the PCO offers meta-classes of a consent management system, the customized DICMO uses the PCO, FOAF, and Relationship Ontology to integrate consent management concepts with the relevant concepts of the related domain.

D. FOAF AND RELATIONSHIP ONTOLOGY

FOAF (Friend-Of-A-Friend) [10] defines people, their activities, and relationships with other people and objects. FOAF identifies people and their friends on networks in a machine-readable format. Therefore, these networks can be analyzed by computers and visualized in a more understandable way. For this purpose, FOAF specifies the ontological representation of personal information. Each FOAF user is responsible for their own FOAF document. The FOAF profile of a person describes profile terms related to a person, such as name, surname, gender, and age. The FOAF documents are online documents and are identified with a unique URI. Thus, FOAF documents can also be accessed over the Internet. FOAF is

used in various domains such as healthcare, education, and social networks. The FOAF document represents the person's information with classes and object properties. However, these classes and properties may be insufficient to meet the requirements of the domain. In such cases, the FOAF ontology is extended by adding new classes and properties. In this study, the necessity of adding new structures has emerged as FOAF classes and features are not sufficient for the consent management process.

In the proposed DICON model, the *Person* class corresponding to the *Person* entity is created within the PCO. Hence, the *Person* class in the PCO is mapped to the *Person* class in the FOAF. Thus, the characteristics of the *Person* class is derived from the FOAF. Also, a FOAF document can be associated with another FOAF document by using the *foaf:knows* object property. The *foaf:knows* relationship represents a connected network between people. The *Person* class in the PCO is also assigned to a person document in the FOAF Ontology with the *foaf:knows* object property. In the consent management process, if the person's consent cannot be obtained, the *foaf:knows* object property is used to reach the related person who is designated and authorized by the person. Besides, in the consent management, the person must be able to give consent in order to obtain consent. For example, the individual whose consent is needed may be under the age of 18 or have a psychiatric disability. In such cases, consent is obtained from the person's parent or legal representative. In order to control this constraint, the FOAF Ontology is extended by defining the *foaf:hasMinAge* object property and the *canGiveConsent* data property. Therefore, the person's condition to give consent is represented with these extensions.

The Relationship Ontology [11] is a vocabulary for describing relationships between people, such as mother, father, and legal representative. In the DICON model, the Relationship Ontology is imported in the FOAF to define the relationship between the person, the parent of the person, and the legal representative of the person. Then, the final FOAF is imported into the DICMO to provide the relevant definitions for obtaining the consent of the person. Therefore, at the instance level, the related person, the consent of the person, and relationships of the person are created.

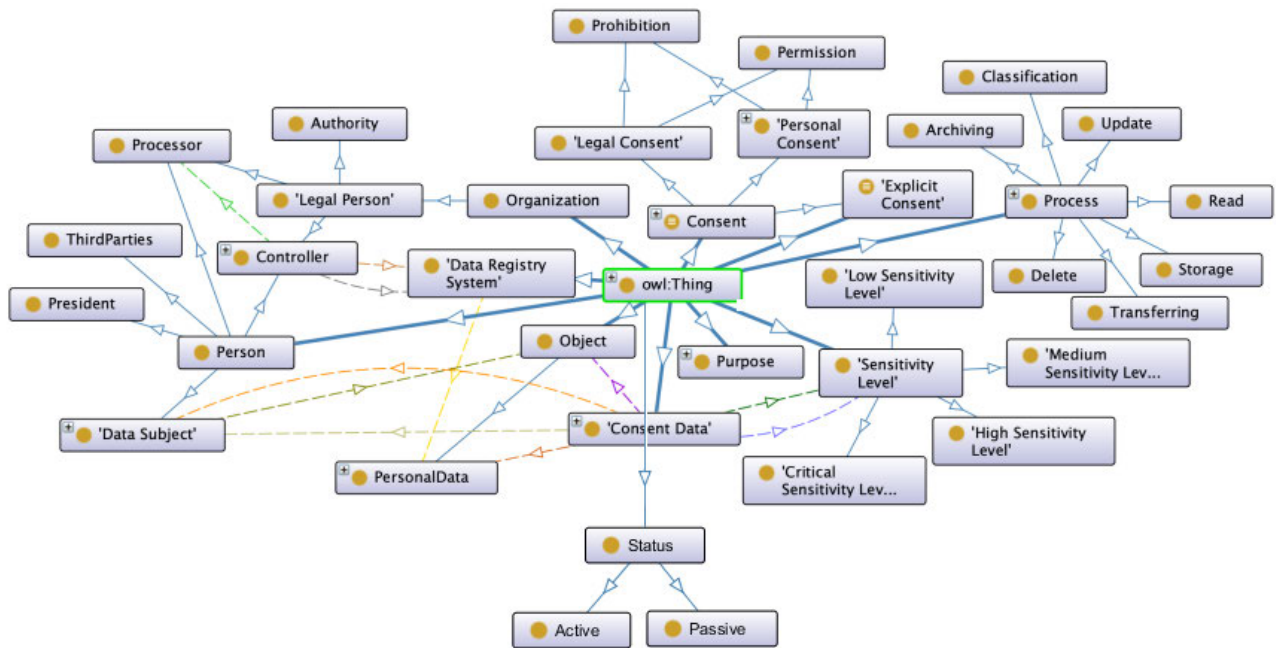


FIGURE 4. The visualization of the PCO.

In this study, the Relationship Ontology is also extended according to the needs of the consent management process. In the Relationship Ontology, there are no relationships such as mother, father, legal representative. After the Relationship Ontology is included in the FOAF Ontology, the `foaf:motherOf`, `foaf:fatherOf`, `foaf:representativeOf` subproperties are defined under the `foaf:knows` object property. Thus, an individual's consent can be obtained from the legal representative or the parent of the individual.

The details of extending FOAF and Relationship Ontology according to the needs of the proposed consent management model can be found in [47].

E. E-Pulse ONTOLOGY

The personalized consent management model proposed in this study is a domain-independent model. In order to evaluate the proposed model, the e-Pulse system which is a trusted personal health record system of Turkey [13] is chosen as the application domain. e-Pulse is a health information system that stores health data that is collected from health institutions and allows citizens and health professionals to access this health data by using the web-based system or the mobile application. The system aims to access personal health data efficiently and effectively through the information systems. The e-Pulse Ontology is based on the e-Pulse system and it is a new ontology that is created as a part of this study. The e-Pulse Ontology stores persons' health data to share them with the necessary institutions and to make inferences from these health data. The subjects and objects of the e-Pulse system are Patient, Physician, Clinic, Family

Medicine, Prescription, Pulse, Body Mass Index, Blood Sugar, Blood Pressure, Health Facility, Hospital Visits, and Hospital Tracking Number. Therefore, the related classes are defined in the e-Pulse Ontology. Also, the Patient class in the e-Pulse Ontology is mapped to the Person class in the FOAF to represent the patient's information in the e-Pulse system. For the evaluation of the proposed model, PurposeOfUse Ontology is also used in addition to the e-Pulse Ontology. the PurposeOfUse Ontology [14] is developed by HL7 to describe the purpose of use for accessing an object in the healthcare domain. Therefore, the PurposeOfUse Ontology is used to represent the access purposes for the use case study.

F. ONTOLOGY EVALUATION

An ontology developed in any ontology development language should be evaluated for knowledge representation before it can be used in an application. Ontology validation tools and ontology platforms are used to evaluate RDF(S), DAML+OIL and OWL ontologies [61]. The goal of ontology evaluation is to determine what the ontology accurately describes and does not describe. Besides, ontology evaluation should be performed throughout the entire ontology lifecycle. In order to evaluate an ontology, ontology parsers are used to identify taxonomic problems in an ontology in terms of information representation.

In this study, W3C RDF Validation Service [62], the SSN Ontology Validation Service [63] and the OOPS! tool [64] are used to evaluate the Personal Consent Ontology. W3C RDF Validation Service and the OOPS! tool provide the

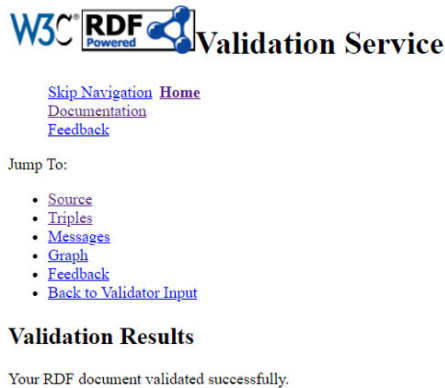


FIGURE 5. RDF Validation of the PCO.

environment for the evaluation. Further, the OOPS! tool presents the evaluation results in a detailed manner.

W3C RDF Validation Service is based on HP-Labs Another RDF Parser (ARP17). The validation service does not offer RDFS validation. However, it offers syntactic validation to check if the developed ontology conforms to the RDF/XML syntax. For the purpose of this study, an RDF/XML document of PCO is given as an input to the RDF Validation Service for the evaluation. After the evaluation, a three-tuple representation of the corresponding data model and a graphical visualization of the data model are presented. The result of the W3C RDF Validation Service concludes that the PCO is valid for RDF as seen in Figure 5.

The PCO is also evaluated with the SSN Ontology Validation Service. An ontology can be either uploaded as a file or entered directly into the textbox to be evaluated with the SSN Ontology Validation Service. The evaluation result is shown in Figure 6. As seen in the figure, one statement is reported as a result of the evaluation. However, this statement is about `owl:qualifiedCardinality` which determines the cardinality of an exact qualified cardinality restriction. This report is not related with the PCO and does not affect the evaluation of the ontology. Therefore, no action needs to be taken for the related result.

Finally, the PCO is evaluated with the OOPS! tool. The evaluation process can be done in two different ways with the OOPS! tool. In the first option, ontology is evaluated according to the criteria that are determined by the OOPS! tool. In the second option, these criteria are grouped and the ontology is evaluated by selecting the related group. In this study, the PCO is evaluated with the first evaluation option which is "Select Pitfalls for Evaluation". The obtained results are as follows:

- Creating unconnected ontology elements
- Missing annotations
- Missing disjointness
- Missing domain or range in properties
- Inverse relationships not explicitly declared
- Defining multiple domain or ranges in properties
- No license declared

After these evaluation results, the PCO is edited and the related errors are corrected. The evaluation results are presented in Figure 7. As seen from the Figure 7, the evaluation results are minor problems that do not affect the technical evaluation of the proposed DICON model.

In the literature, there are different evaluation approaches for the ontology evaluation, such as OntoClean [65] and OntoQA [66]. OntoQA evaluates ontologies based on vocabularies and examples defined in an OWL/RDFs document. For this purpose, OntoQA focuses on the arrangement of classes in the schema and the distribution of examples throughout the schema. The metrics used in OntoQA are divided into two categories: Schema Metrics and Instance Metrics. While Schema Metrics evaluate the design of the ontology, Instance Metrics evaluate the placement of instance data within the ontology and the effective use of the ontology to represent the information modeled in the ontology. For the ontology design, it may not be clearly known whether the information is modeled appropriately or not. However, the design evaluation of an ontology can be addressed by providing metrics such as richness, width, depth and inheritance of an ontology schema. The Instance Metrics are focused on how data is placed in an ontology. The layout of the data is an important criterion for the quality of an ontology. The effectiveness of the ontology design can be demonstrated by the placement of the sample data and the distribution of the data. The Instance Metrics are divided into two categories, Knowledge Base (KB) metrics and class metrics that describe how each class is used in the KB.

In this study, the evaluation of PCO with Class Importance under Instance Metrics is also presented. Class Importance is calculated with the percentage of instances belonging to classes in the subtree rooted in the current class relative to the total number of instances. The related metric is also called the Sample Distribution. The Class Importance is an important metric as it allows to determine which fields of the schema are in focus when samples are extracted from the ontology model. Besides, the related metric informs the user on the appropriateness of the schema's intended use.

The formula for the Class Importance metric is given in Equation (2). In the Equation (2), the importance of Class (C_i) is denoted by Imp . The Imp of the C_i class is defined by comparing the number of Instances (I) of the subtree rooted in C_i in KB ($C_i(I)$) with the total number of instances in KB (I).

$$Imp = \frac{|C_i(I)|}{|I|} \quad (2)$$

According to the results of calculations given in Figure 8, the Person, Object, and Consent classes are the dominant classes of the PCO. Except for the Registry of Controllers class, it gives consistent importance to most of the classes it contains. Since the number of subclasses under Process is high, the number of instances is also high. The prominent classes of GDPR ontology [67] are also calculated and presented in Figure 9. According to Figure 9, the GDPR ontology mostly includes legislative classes than a

Validation Report

On statement: `_:b1001 owl:qualifiedCardinality "1"^^http://www.w3.org/2001/XMLSchema#nonNegativeInteger`
 predicate not declared in any schema: `owl:qualifiedCardinality`

FIGURE 6. The evaluation result with SSN Ontology Validation Service.

Evaluation results

It is obvious that not all the pitfalls are equally important; their impact in the ontology will depend on multiple factors. For this reason, each pitfall has an importance level attached indicating how important it is. We have identified three levels:

- **Critical** 🚫 : It is crucial to correct the pitfall. Otherwise, it could affect the ontology consistency, reasoning, applicability, etc.
- **Important** ⚠️ : Though not critical for ontology function, it is important to correct this type of pitfall.
- **Minor** 🟡 : It is not really a problem, but by correcting it we will make the ontology nicer.

[Expand All] | [Collapse All]

| | |
|---|--------------------------|
| Results for P02: Creating synonyms as classes. | 1 case Minor 🟡 |
| Results for P04: Creating unconnected ontology elements. | 2 cases Minor 🟡 |
| Results for P13: Inverse relationships not explicitly declared. | 13 cases Minor 🟡 |
| Results for P41: No license declared. | ontology* Important ⚠️ |

FIGURE 7. The evaluation result with OOPS! tool.

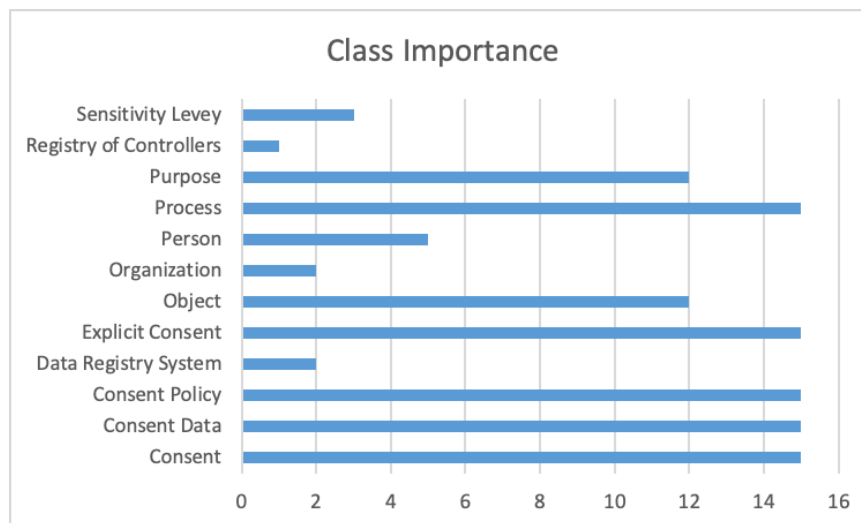


FIGURE 8. The Class Importance of the PCO.

structure that practically supports the consent management process. Consequently, the Class Importance metric is used to compare the proposed PCO with the GDPR ontology. The results clearly show their differences and their purposes of use.

IV. CONSENT MANAGEMENT FOR PRIVACY PROTECTION: A USE CASE FOR HEALTHCARE

Today, informed consent is mandated for almost every system with the enforcement of regulations and laws. In this study, the healthcare domain is chosen as the application domain for the proposed consent management model. The SWRL [70]

rule definitions are given in Table 2. According to these definitions, Rule 1 and Rule 2 check the *Data Subject's* ability to consent by looking at her age (LPPD states that if a person's age is over 18 she can give consent, in GDPR this age limit is 16); Rule 3 states that there is no need for parental control depending on the *Data Subject's* age; Rule 4 indicates that parental control is required depending on the age of the person; Rule 5 checks whether the *Institution* which carries out the process for the protection of *Personal Data* has *Permission* to transfer personal data abroad or not. Moreover, Rule 6 and Rule 7 checks the sensitivity levels for the consent decision. The SWRL rule validation,

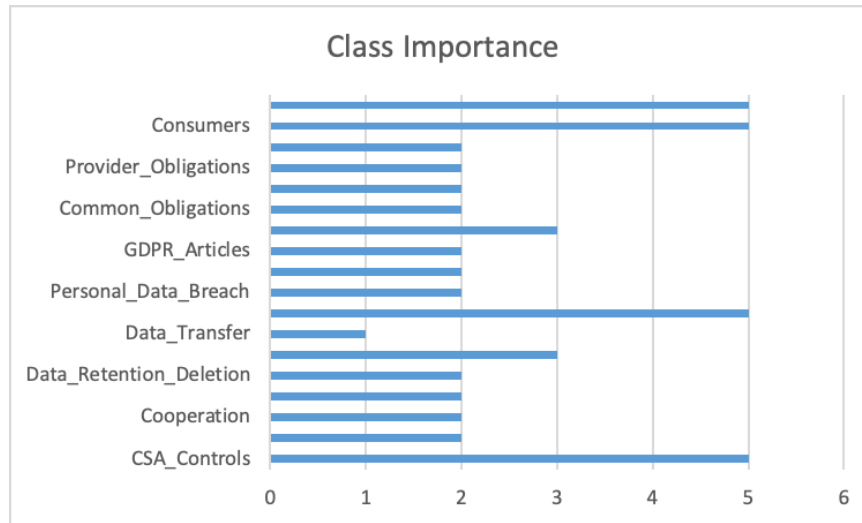


FIGURE 9. The Class Importance of the GDPR Ontology.

TABLE 2. SWRL rules for the consent management model.

| |
|--|
| Rule-1: Person is capable of giving consent |
| <code>foaf:canGiveConsent(?p,true) ← PersonalConsentOntology:DataSubject(?p) ∧ fco:age(?p,?age) ∧ swrlb:greaterThanOrEqual(?age,18)</code> |
| Rule-2: Person is not capable of giving consent |
| <code>foaf:canGiveConsent(?p,false) ← PersonalConsentOntology:DataSubject(?p) ∧ ePulse:age(?p,?age) ∧ swrlb:lessThan(?age,18)</code> |
| Rule-3: No need for parental consent. |
| <code>PersonalConsentOntology:parentalControl(?p,"No need for parental consent.") ← PersonalConsentOntology:DataSubject(?p) ∧ fco:age(?p,?age) ∧ swrlb:greaterThanOrEqual(?age,18)</code> |
| Rule-4: Parental consent must be obtained. |
| <code>PersonalConsentOntology:parentalControl(?p,"Parental consent must be obtained.") ← PersonalConsentOntology:DataSubject(?p) ∧ fco:age(?p,?age) ∧ swrlb:lessThan(?age,18)</code> |
| Rule-5: The institution does not allow to share information abroad. |
| <code>PersonalConsentOntology:hasConsentType(?policy, dicmo:prohibition) ← PersonalConsentOntology:hasProcess(?policy,?i) ∧ PersonalConsentOntology:abroad(?i,false)</code> |
| Rule-6: Give warning if the Sensitivity level is "Critical". |
| <code>PersonalConsentOntology:giveWarningFor(?cd,true) ← PersonalConsentOntology:DataSubject(?p) ∧ PersonalConsentOntology:holdSubject(?cd,?p) ∧ PersonalConsentOntology:hasSensitivityLevel(?cd,'Critical') ∧ PersonalConsentOntology:hasOtherConsent(?cd,true)</code> |
| Rule-7: Checking Sensitivity level for consent policy. |
| <code>PersonalConsentOntology:hasConsentType(?policy,ePulse:?c ← PersonalConsentOntology:DataSubject(?p) ∧ PersonalConsentOntology:holdSubject(?cd,?p) ∧ PersonalConsentOntology:hasOtherConsent(?cd,true) ∧ (PersonalConsentOntology:hasSensitivityLevel(?cd,'Low') PersonalConsentOntology:hasSensitivityLevel(?cd,'Medium'))</code> |

the instances after the validation related with Rule 1 and Rule 2 are shown in Figure 10, Figure 11 and Figure 12, respectively.

In order to evaluate the domain-independent consent management model, a Personalized Consent Management System is developed based on the proposed model. The system is implemented by using Apache Jena [68]. The implemented

consent management system enables to enforce the consent management process on a person's pulse information. For this purpose, e-Pulse Ontology is used as the domain ontology.

The scenario created within the context of the case study is as follows: Patients' health data is kept electronically by the Ministry of Health. The Ministry of Health stores the health data in its database and it is the responsible

| Active ontology × Entities × Individuals by class × OWLViz × DL Query × SWRLTab × | | |
|---|-----------------------------------|--|
| | Name | Rule |
| <input checked="" type="checkbox"/> | capableOfGivingConsent | pdplontology:DataSubject(?p) ^ fco:age(?p, ?age) ^ swrlb:greaterThanOrEqual(?age, 18) -> autogen0:canGiveConsent(?p, true) |
| <input checked="" type="checkbox"/> | institutionDoesNotAllowToShare... | pdplontology:hasProcess(?policy, ?i) ^ pdplontology:abroad(?i, false) -> pdplontology:hasConsentType(?policy, dicmo:prohibition) |
| <input checked="" type="checkbox"/> | noNeedForParentalConsent | pdplontology:DataSubject(?p) ^ fco:age(?p, ?age) ^ swrlb:greaterThanOrEqual(?age, 18) -> pdplontology:parentalControl(?p, "No need for parental consent.") |
| <input checked="" type="checkbox"/> | notCapableOfGivingConsent | pdplontology:DataSubject(?p) ^ fco:age(?p, ?age) ^ swrlb:lessThan(?age, 18) -> autogen0:canGiveConsent(?p, false) |
| <input checked="" type="checkbox"/> | parentalConsentMustBeObtained | pdplontology:DataSubject(?p) ^ fco:age(?p, ?age) ^ swrlb:lessThan(?age, 18) -> pdplontology:parentalControl(?p, "No need for parental consent.") |

| | | | | |
|---------|-------|-----------------|-----------------|----------|
| Control | Rules | Asserted Axioms | Inferred Axioms | OWL 2 RL |
|---------|-------|-----------------|-----------------|----------|

OWL axioms successfully transferred to rule engine.
 Number of SWRL rules exported to rule engine: 5
 Number of OWL class declarations exported to rule engine: 89
 Number of OWL individual declarations exported to rule engine: 3
 Number of OWL object property declarations exported to rule engine: 66
 Number of OWL data property declarations exported to rule engine: 34
 Total number of OWL axioms exported to rule engine: 970
 The transfer took 89 millisecond(s).
 Press the 'Run Drools' button to run the rule engine.
 Successful execution of rule engine.
 Number of inferred axioms: 737
 The process took 118 millisecond(s).
 Look at the 'Inferred Axioms' tab to see the inferred axioms.
 Press the 'Drools->OWL' button to translate the inferred axioms to OWL knowledge.
 Successfully transferred inferred axioms to OWL model.
 The process took 17 millisecond(s).

FIGURE 10. The SWRL validation.

Annotations Usage

Usage: dicmo:Lisa

Show: ☒ this ☒ different

Found 30 uses of dicmo:Lisa

dicmo:Lisa

- dicmo:Lisa foaf:canGiveConsent true
- dicmo:Lisa Type pdplontology:Person
- dicmo:Lisa Type schema:Person
- dicmo:Lisa fco:age "37"^^xsd:int
- dicmo:Lisa Type foaf:OWLClassImpl_5d12512b_448e_480a_bd2e_f477772a1cbf
- dicmo:Lisa Type dcterms:Agent
- dicmo:Lisa Type geo:SpatialThing
- dicmo:Lisa foaf:Mother_Of dicmo:mary
- dicmo:Lisa Type foaf:OWLClassImpl_be1109f3_620f_490b_85ed_c14c0ab82e84
- dicmo:Lisa owl:differentFrom dicmo:mary
- dicmo:Lisa Type <http://www.w3.org/2000/10/swap/pim/contact#Person>
- dicmo:Lisa foaf:OWLObjectPropertyImpl_e5e5135c_aa01_46d6_8904_2882ac426498 dicmo:mary

Description: dicmo:Lisa

Types

- <http://www.w3.org/2000/10/swap/pim/contact#Person>
- dcterms:Agent
- foaf:OWLClassImpl_5d12512b_448e_480a_bd2e_f477772a1cbf
- foaf:OWLClassImpl_be1109f3_620f_490b_85ed_c14c0ab82e84
- geo:SpatialThing
- pdplontology:DataSubject
- pdplontology:Organization
- pdplontology:Person
- schema:Person

Property assertions: dicmo:Lisa

Object property assertions

- foaf:Mother_Of (foaf:Mother_Of) dicmo:mary
- owl:differentFrom dicmo:mary
- foaf:OWLObjectPropertyImpl_e5e5135c_aa01_46d6_8904_2882ac426498 dicmo:mary

Data property assertions

- foaf:canGiveConsent true
- fco:age "37"^^xsd:int

Negative object property assertions

FIGURE 11. Instance for Rule-1.

institution of this database. The Public Health Unit, one of the units of the Ministry of Health, wants to classify the e-Pulse data for research purposes. Jack Smith is a patient and he splits his Personal Data as DemographicData and PersonalPulseData. Jack Smith determines two different consent for his DemographicData and PersonalPulseData. The related atomic concepts and roles of the case study are given in *ALCQ* DL as follows:

$\text{hasDataRegistrySystem}(\text{MinistryofHealth}) = \text{ePulseDatabase}$
 $\text{responsibleFor}(\text{MinistryofHealth}) = \text{ePulseDatabase}$
 $\text{registered}(\text{MinistryofHealth}) = \text{registryOfDataControllers}$

$\text{authorize}(\text{MinistryofHealth}) = \text{PublicHealthUnit}$
 $\text{hold}(\text{ePulseDatabase}) = \text{JackSmithPersonalPulseData}$
 $\text{hold}(\text{ePulseDatabase}) = \text{JackSmithDemographicData}$
 $\text{hasObject}(\text{JackSmith}) = \text{JackSmithPersonalPulseData}$
 $\text{hasObject}(\text{JackSmith}) = \text{JackSmithDemographicData}$
 $\text{hasPulse}(\text{JackSmithPersonalPulseData}) = 67$
 $\text{hasBloodSugar}(\text{JackSmithPersonalPulseData}) = 87$
 $\text{hasBloodPressure}(\text{JackSmithPersonalPulseData}) = 87$

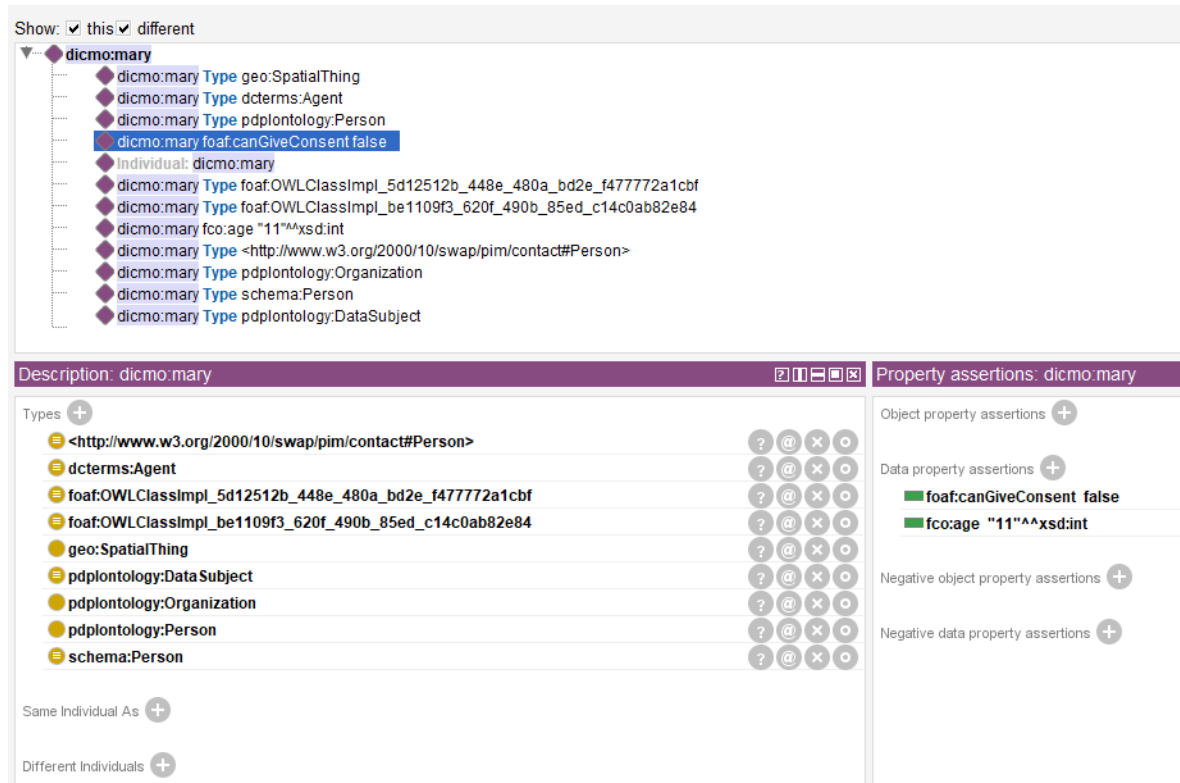


FIGURE 12. Instance for Rule-2.

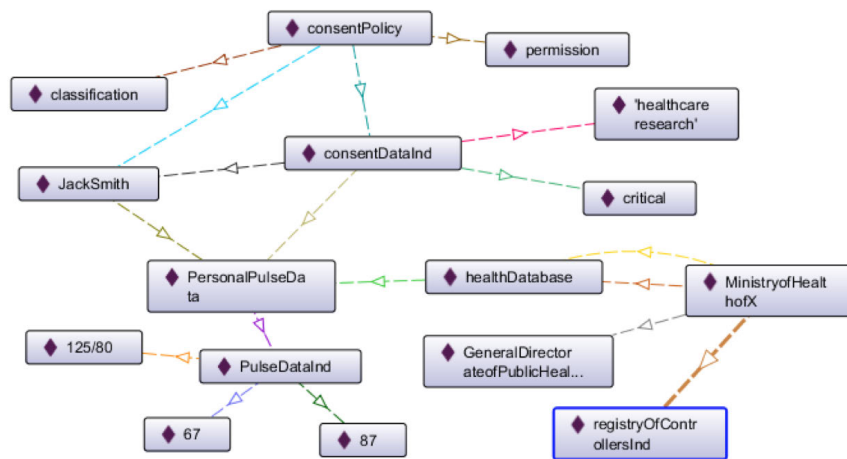


FIGURE 13. The ontology visualization of the use case study.

Data) = 125/80
 Jack Smith's consent definitions for his data are given in below:
 holdSubject (consentData1) = JackSmith
 holdObject (consentData1) =
 JackSmithPersonalPulseData
 hasSensitivityLevel (consentData1) =
 critical
 holdSubject (consentData2) = JackSmith

holdObject (consentData2) =
 JackSmithDemographicData
 hasSensitivityLevel (consentData2) = high
 Jack Smith's consent policy definitions are as follows:
 hasConsentData (consentPolicy1) =
 consentData1
 hasConsentType (consentPolicy1) =
 permission
 hasConsenter (consentPolicy1) = JackSmith

| SPARQL query: | | | | |
|--|-------------------|----------------|-----------------------|-------------|
| <pre> PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX owl: <http://www.w3.org/2002/07/owl#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> PREFIX xsd: <http://www.w3.org/2001/XMLSchema#> PREFIX pco: <http://www.semanticweb.org/user/ontologies/2021/2/PCOntology.owl#> PREFIX epulse: <http://www.semanticweb.org/user/ontologies/2020/2/ePulseOntology#> PREFIX relationship: <http://purl.org/vocab/relationship/> PREFIX foaf: <http://xmlns.com/foaf/0.1/> SELECT ?DataSubject ?PersonalData ?Process ?Purpose ?Consent WHERE { ?DataSubject foaf:givenName "JackSmith"^^xsd:string. ?conPol pco:hasConsenter ?DataSubject . ?conPol pco:hasConsentData ?ConData . ?conPol pco:hasProcess ?Process . ?conPol pco:hasPurposeofUse ?Purpose . ?ConData pco:holdObject ?PersonalData . ?conPol pco:hasConsentType ?Consent } </pre> | | | | |
| DataSubject | PersonalData | Process | Purpose | Consent |
| JackSmith | DemographicData | classification | 'healthcare research' | prohibition |
| JackSmith | PersonalPulseData | classification | 'healthcare research' | permission |

FIGURE 14. The SPARQL query and the result of the query for Jack Smith's consents.

```

hasProcess (consentPolicy1) =
  classification
hasPurposeofUse (consentPolicy1) = HRESCH
hasConsentData (consentPolicy2) =
  consentData2
hasConsentType (consentPolicy2) =
  prohibition
hasConsenter (consentPolicy2) = JackSmith
hasProcess (consentPolicy2) =
  classification
hasPurposeofUse (consentPolicy2) = HRESCH
hasConsenter (consentPolicy1, JackSmith) ==
  hasConsentPolicy (JackSmith,
    consentPolicy1)
hasPermission (consentPolicy1) =
  (JackSmith, JackSmithPersonalPulseData,
    classification, HRESCH, permission)
hasConsenter (consentPolicy2, JackSmith) ==
  hasConsentPolicy (JackSmith,
    consentPolicy2)
hasPermission (consentPolicy2) =
  (JackSmith, JackSmithDemographicData,
    classification, HRESCH, prohibition)

```

The visualization of the presented use case study is shown in Figure 13. In Figure 14, a SPARQL query [69] and its results are given. In the given query, Jack Smith's personal data, permissions, prohibitions, access purposes and type of data process on his personal data are queried. The result of this query shows Jack Smith's consents, his personal data sets, the type of process to be performed on his personal data and the purpose of the related process. As seen in the query result, the first consent states that accessing to DemographicData to perform the classification process for research purposes is prohibited and the second consent states that accessing to PersonalPulseData to perform the classification process for research purposes is permitted.

Figure 15 shows the query result of Jack Smith's personal data that is performed on the DICON system.

FIGURE 15. Jack Smith's personal data.

FIGURE 16. Jack Smith's consent on his pulse data.

As seen in the figure, Jack Smith's pulse data and demographic data are listed under his Personal Data. Jack Smith's consents related with his personal data are given in Figure 16 and Figure 17, respectively. Figure 16 states that Jack Smith's pulse data has a Critical sensitivity level

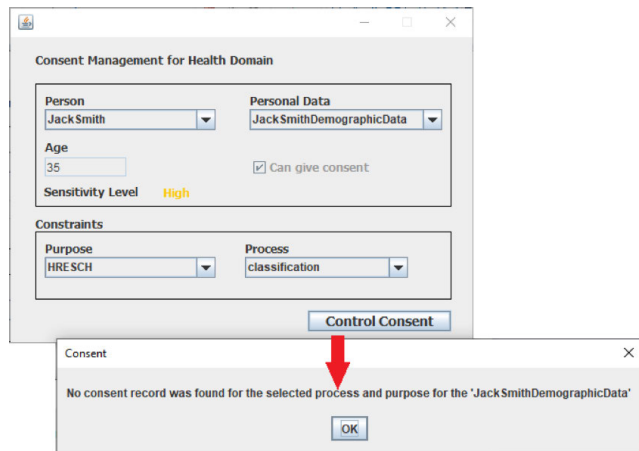


FIGURE 17. Jack Smith's consent on his demographic data.

and he has consent for healthcare research purposes which is stated as HRESCH. However, Figure 17 shows that Jack Smith has a High sensitivity level for his demographic data and he has no consent for this data to be used with the same purpose.

V. CONCLUSION AND FUTURE WORK

Collecting, processing, and revealing personal data threaten privacy and data security. Consent management aims to protect privacy by using individuals' consent information on their personal data. In this study, a Semantic Web-based domain-independent consent management approach is proposed. The proposed model is based on the Law on Protection of Personal Data (LPPD) of Turkey. According to this Law, the consent of an individual is necessary for the usage of personal data in every system where personal data is stored. Therefore, the functional elements of the proposed model are extracted from the LPPD. In this study, a well-defined ontological model of the consent management process is proposed. For this purpose, the proposed model's fundamental elements, actors, and relationships between them are presented. Also, ontologies that are created for the model are presented. As the core ontologies of the model are domain-independent, the model can be integrated into any domain. In order to evaluate the proposed domain-independent consent management model, the implementation is performed for the healthcare domain. The e-Pulse Ontology that is created for the use case study of the proposed model is the first ontology that is based on the national e-Pulse personal health record system of Turkey.

The presented domain-independent model is the first ontology-based consent management model that is based on Personal Data Protection Law. Ontologies that are unique to the presented consent management model are integrated with FOAF and Relationship Ontology which are extended according to the needs of the consent management process. Therefore, the model also provides a personalized privacy

approach to preserve privacy and to support personalized privacy.

As future work, e-Pulse Ontology will be extended, new individuals will be added and different SPARQL queries will be executed. Besides the healthcare domain, the domain-independent model will also be evaluated for the education domain. Also, a conflict engine will be implemented to determine and resolve consent conflicts that may occur. Moreover, it will be ensured that the personalized consent management system also detects possible privacy violations and prevents data breaches that may happen during information processing.

REFERENCES

- [1] EU's General Data Protection Regulation (GDPR). 2016/679 Of the European Parliament and of The Council of April 16, 2016. Accessed: Jul. 10, 2022. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>
- [2] N. Gruschka and M. Jensen, "Aligning user consent management and service process modeling," in *Proc. Informatik*. Stuttgart, Germany, Sep. 2014, pp. 527–538.
- [3] H. J. Pandit, C. Debruyne, D. O'Sullivan, and D. Lewis, "GConsent—A consent ontology based on the GDPR," in *Proc. 16th Int. Eur. Semantic Web Conf. (ESWC)* in Lecture Notes in Computer Science, vol. 11503, Portorož, Slovenia: Springer, Jun. 2019, pp. 270–282.
- [4] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Sci. Amer.*, vol. 284, no. 5, pp. 34–43, 2001.
- [5] M. M. Taye, "Understanding semantic web and ontologies: Theory and applications," *J. Comput.*, vol. 2, no. 6, pp. 182–192, 2010.
- [6] T. R. Gruber, "A translation approach to portable ontologies," *Knowl. Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.
- [7] UNCTAD. (2020). *Data Protection and Privacy Legislation Worldwide*. Accessed: Jul. 10, 2022. [Online]. Available: <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide>
- [8] Personal Data Protection Authority. (2016). *Law on The Protection of Personal Data*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.kvkk.gov.tr/SharedFolderServer/CMSFiles/aea97a33-089b-4e7d-85cb-694adb57bed3.pdf>
- [9] Koyuncuoglu & Koksall Law Firm. (2018). *What Are the Differences Between the Data Protection Law and GDPR*. Accessed: Jul. 10, 2022. [Online]. Available: <https://kkhukuk.com/wp-content/uploads/2018/07/What-are-the-differences-between-the-Data-Protection-Law-and-GDPR-28072018.pdf>
- [10] D. Brickley and L. Miller. (2014). *FOAF Vocabulary Specification 0.99*. Accessed: Jul. 10, 2022. [Online]. Available: <http://xmlns.com/foaf/spec>
- [11] I. Davis. (2004). *Relationship: A Vocabulary for Describing Relationships Between People*. Accessed: Jul. 10, 2022. [Online]. Available: <http://vocab.org/relationship>
- [12] A. F. Westin, *Privacy Freedom*. New York, NY, USA: Atheneum, 1967.
- [13] T. C. Ministry of Health. *e-Pulse (e-Nabiz) Personal Health System*. Accessed: Jul. 10, 2022. [Online]. Available: <https://enabiz.gov.tr>
- [14] HL7 Security Work Group. *Security and Privacy Ontology*. Accessed: Jul. 10, 2022. [Online]. Available: https://wiki.hl7.org/index.php?title=Security_and_Privacy_Ontology
- [15] R. Clarke, "E-consent: A critical element of trust in e-business," in *Proc. 15th Bled Electron. Commerce Conf. eReality, Constructing eEconomy*, vol. 19. Bled, Slovenia, Jun. 2002, pp. 338–360.
- [16] (1957). *Salgo V. Leland Stanford Etc. Bd. Trustees, 154 Cal.App.2d 560*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.leagle.com/decision/1957714154calapp2d5601626>
- [17] AMA (American Medical Association). (2008). *Informed Consent*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.ama-assn.org/delivering-care/informed-consent>
- [18] J. B. Williams and T. Figley, "System and method for providing consent management," WO Patent 2016060858, Apr. 11, 2016.
- [19] (1996). *Health Insurance Portability and Accountability Act (HIPAA)*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.hhs.gov/hipaa/index.html>

- [20] (1999). *Gramm-Leach-Bliley Act (GLBA)*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.ftc.gov/tips-advice/business-center/privacy-and-security/gramm-leach-bliley-act>
- [21] J. C. Vargas, "Blockchain-based consent manager for GDPR compliance," in *Proc. Open Identity Summit in Lecture Notes in Informatics (LNI)*. Bonn, Germany: Gesellschaft Für Informatik, 2019, pp. 165–170.
- [22] S. Tokas and O. Owe, "A formal framework for consent management," in *Proc. Int. Conf. Formal Techn. Distrib. Objects, Compon., Syst. (FORTE)*, vol. 12136. Valletta, Malta, Jun. 2020, pp. 169–186.
- [23] I. Horrocks, "Ontologies and the semantic web," *Commun. ACM*, vol. 51, no. 12, pp. 58–67, 2008.
- [24] O. Can, "A semantic model for personal consent management," in *Proc. 7th Int. Metadata Semantics Res. Conf. (MTSR)*, vol. 390. Thessaloniki, Greece, Nov. 2013, pp. 146–151.
- [25] E. Olca and O. Can, "A meta-consent model for personalized data privacy," presented at the 10th Int. Metadata Semantics Res. Conf. (MTSR), Göttingen, Germany, Nov. 2016.
- [26] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini, and L. Robaldo, "Legal ontology for modelling GDPR concepts and norms," in *Legal Knowledge and Information Systems*, vol. 313. Amsterdam, The Netherlands: IOS Press, 2018, p. 91100.
- [27] M. Laurent, J. Leneutre, S. Chabridon, and I. Laouane, "Authenticated and privacy-preserving consent management in the Internet of Things," *Proc. Comput. Sci.*, vol. 151, pp. 256–263, Jan. 2019.
- [28] E. Luger and T. Rodden, "Terms of agreement: Rethinking consent for pervasive computing," *Interacting With Comput.*, vol. 25, no. 3, pp. 229–241, May 2013.
- [29] E. Coiera and R. Clarke, "e-Consent: The design and implementation of consumer consent mechanisms in an electronic environment," *J. Amer. Med. Inform. Assoc.*, vol. 11, no. 2, pp. 40–129, 2004.
- [30] J. Hyysalo, H. Hirvonsalo, J. Sauvola, and S. Tuoriniemi, "Consent management architecture for secure data transactions," in *Proc. 11th Int. Joint Conf. Softw. Technol.*, Jul. 2016, pp. 125–132.
- [31] K. E. Berntsen and V. Heimly, "Consent-based access to core EHR information: Collaborative approaches in Norway," *Methods Inf. Med.*, vol. 48, no. 2, pp. 144–148, 2009.
- [32] O. Heinze, M. Birkle, L. Köster, and B. Bergh, "Architecture of a consent management suite and integration into IHE-based regional health information networks," *BMC Med. Informat. Decis. Making*, vol. 11, no. 1, p. 58, Dec. 2011.
- [33] *Electronic Consent Management: Landscape Assessment, Challenges, and Technology*, Version 1.0., MITRE Corporation, McLean, VA, USA, 2014.
- [34] N. P. Sheppard, R. Safavi-Naini, and M. Jafari, "A digital rights management model for healthcare," in *Proc. IEEE Int. Symp. Policies Distrib. Syst. Netw.*, Jul. 2009, pp. 106–109.
- [35] M. A. Grando, A. Boxwala, R. Schwab, and N. Alipanah, "Ontological approach for the management of informed consent permissions," in *Proc. IEEE 2nd Int. Conf. Healthcare Informat., Imag. Syst. Biol.*, Sep. 2012, pp. 51–60.
- [36] B. Yu, D. Wijesekera, and P. Costa, "An ontology for medical treatment consent," in *Proc. Semantic Technol. Intell., Defense, Secur. (STIDS)*, vol. 1304. Fairfax, VA, USA, Nov. 2014, pp. 72–79.
- [37] Y. Lin, M. R. Harris, F. J. Manion, E. Eisenhauer, B. Zhao, W. Shi, A. Karnovsky, and Y. He, "Development of a BFO-based informed consent ontology (ICO)," in *Proc. 5th Int. Conf. Biomed. Ontology (ICBO)*, vol. 1327. Houston, TX, USA, Oct. 2014, pp. 84–86.
- [38] Y. Lin, J. Zheng, and Y. He, "VICO: Ontology-based representation and integrative analysis of vaccination informed consent forms," *J. Biomed. Semantics*, vol. 7, no. 1, p. 20, Dec. 2016.
- [39] HL7. (2014). *HL7 Version 3 Standard: Security and Privacy Ontology, Release 1*. Accessed: Jul. 10, 2022. [Online]. Available: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=348
- [40] K. P. Joshi. (2018). *Ontology for EU's General Data Protection Regulation (GDPR)*. UMBC ebiquity Research Group. Accessed: Jul. 10, 2022. [Online]. Available: <https://ebiquity.umbc.edu/resource/html/id/377/Ontology-for-EU-s-General-Data-Protection-Regulation-GDPR>
- [41] HL7 Workgroup. (2010). *Security and Privacy Ontology*. Accessed: Jul. 10, 2022. [Online]. Available: http://wiki.hl7.org/index.php?title=Security_and_Privacy_Ontology
- [42] HL7 International. (2007). *Introduction to HL7 Standards*. Accessed: Jul. 10, 2022. [Online]. Available: <http://www.hl7.org/implement/standards/index.cfm?ref=nav>
- [43] *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*, OECD, Paris, France, 2013.
- [44] U. Kilkelly, "The right to respect for private and family life—A guide to the implementation of article 8 of the European convention on human rights," in *Human Rights Handbooks*, no. 1, 2001.
- [45] "Charter of fundamental rights of the European union (2000/C 364/01)," *Off. J. Eur. Communities*, 2000. Accessed: Sep. 7, 2022. [Online]. Available: https://www.europarl.europa.eu/charter/pdf/text_en.pdf
- [46] *Convention for the Protection of Individuals With Regard to Automatic Processing of Personal Data ETS No:108*, Council of Europe, London, U.K., 1985.
- [47] E. Olca and O. Can, "Extending FOAF and relationship ontologies with consent ontology," in *Proc. 3rd Int. Conf. Comput. Sci. Eng. (UBMK)*, Sep. 2018, pp. 542–546.
- [48] Object Management Group. (1997). *Meta-Object Facility (MOF)*. Accessed: Jul. 10, 2022. [Online]. Available: <http://www.omg.org/mof>
- [49] W. Tang, "Meta object facility," in *Encyclopedia of Database Systems*. Boston, MA, USA: Springer, 2009.
- [50] E. Seidewitz, "What models mean," *IEEE Softw.*, vol. 20, no. 5, pp. 26–32, Sep. 2003.
- [51] *Protégé*. Accessed: Jul. 10, 2022. [Online]. Available: <https://protege.stanford.edu>
- [52] O. Can and E. Olca, "Development of a consent ontology for the law on the protection of personal data," *DEU FMD*, vol. 21, no. 62, pp. 559–575, 2019.
- [53] H. C. Wu, R. W. P. Luk, K. F. Wong, and K. L. Kwok, "Interpreting TF-IDF term weights as making relevance decisions," *ACM Trans. Inf. Syst.*, vol. 26, no. 3, pp. 1–37, Jun. 2008.
- [54] K. T. Frantzi, S. Ananiadou, and J. Tsujii, "The C-value/NC-value method of automatic recognition for multi-word terms," in *Proc. 2nd Eur. Conf. Res. Adv. Technol. Digit. Libraries (ECDL)*, vol. 1513. Crete, Greece: Heraklion, 1998, pp. 585–604.
- [55] A. Kiryakov, B. Popov, I. Terziev, D. Manov, and D. Ognyanoff, "Semantic annotation, indexing, and retrieval," in *Proc. Int. Semantic Web Conf. (ISWC)*, vol. 2870. Sanibel Island, FL, USA, 2003, pp. 484–499.
- [56] R. Navigli and P. Velardi, "From glossaries to ontologies: Extracting semantic structure from textual definitions," in *Proc. Conf. Ontol. Learn. Population, Bridging Gap Between Text Knowl.*, vol. 167. Amsterdam, The Netherlands: IOS Press, 2008, pp. 71–87.
- [57] D. Jurafsky and J. H. Martin. (2021). *Syntactic Parsing*. Chapter 11, Speech and Language Processing. Accessed: Aug. 16, 2022. [Online]. Available: <https://web.stanford.edu/~jurafsky/slp3/11.pdf>
- [58] P. Cimiano and J. Völker, "Text2Onto: A framework for ontology learning and data-driven change discovery," in *Proc. 10th Int. Conf. Appl. Natural Lang. Inf. Syst. (NLDB)*, vol. 3513. Alicante, Spain, 2005, pp. 227–238.
- [59] S. Jupp and M. Horridge. (2008). *TerMine Plugin*. Accessed: Aug. 16, 2022. [Online]. Available: https://protegewiki.stanford.edu/wiki/TerMine_Plugin
- [60] B. Popov, A. Kiryakov, A. Kirilov, D. Manov, D. Ognyanoff, and M. Goranov, "KIM-semantic annotation platform," in *Proc. Int. Semantic Web Conf. (ISWC)*, vol. 2870. Sanibel Island, FL, USA, 2003, pp. 834–849.
- [61] O. Corcho, A. Gómez-Pérez, R. González-Cabero, and M. C. Suárez-Figueroa, "ODEval: A tool for evaluating RDF(S), DAML+OIL, and OWL concept," in *Proc. Artif. Intell. Appl. Innov. (AIAI), IFIP Int. Fed. Inf. Process.*, vol. 154. Boston, MA, USA: Springer, 2004, pp. 369–382.
- [62] E. Prud'hommeaux. (2006). *Validation Service*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.w3.org/RDF/Validator>
- [63] S. Kolozali, "A validation tool for the W3C SSN ontology based sensory semantic knowledge," in *Proc. 6th Int. Workshop Found., Technol. Appl. Geospatial Web 7th Int. Workshop Semantic Sensor Netw.*, vol. 1401, 2014, pp. 83–88.
- [64] M. Poveda. (2021). *OOPS! (Ontology Pitfall Scanner!)*. Accessed: Jul. 10, 2022. [Online]. Available: <https://oops.linkeddata.es>
- [65] N. Guarino and C. Welty, "Evaluating ontological decisions with OntoClean," *Commun. ACM*, vol. 45, no. 2, pp. 61–65, Feb. 2002.
- [66] S. Tartir, I. Budak Arpinar, M. Moore, A. P. Sheth, and B. Aleman-Meza, "OntoQA: Metric-based ontology quality analysis," in *Proc. IEEE ICDM Workshop Knowl. Acquisition From Distributed, Auton., Semantically Heterogeneous Data and Knowledge Sources*, 2005, pp. 1–10.
- [67] K. P. Joshi. (2018). *Ontology for EU's General Data Protection Regulation (GDPR)*. Accessed: Aug. 16, 2022. [Online]. Available: <https://ebiquity.umbc.edu/resource/html/id/377/Ontology-for-EU-s-General-Data-Protection-Regulation-GDPR>
- [68] *Apache Jena*. Accessed: Jul. 10, 2022. [Online]. Available: <https://jena.apache.org>

- [69] W3C Recommendation. (2008). *SPARQL Query Language for RDF*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.w3.org/TR/rdf-sparql-query>
- [70] I. Horrocks, P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosz, and M. Dean. *SWRL: A Semantic Web Rule Language Combining OWL and RuleML*. Accessed: Jul. 10, 2022. [Online]. Available: <https://www.w3.org/Submission/SWRL>



EMRE OLCA received the B.S. degree in computer engineering from Mart University, Canakkale, Turkey, the M.S. degree from the Department of Computer Engineering, Eylul University, Izmir, Turkey, and the Ph.D. degree in computer engineering from the Department of Computer Engineering, Ege University, Izmir.

She has worked as a computer engineer in private sector. From 2013 to 2016, she was a Research Assistant. She is currently working as an Assistant

Professor at the Department of Software Engineering, Maltepe University, Istanbul, Turkey. Her research interests include privacy, knowledge engineering, Semantic Web technologies, and robot programming.



OZGU CAN received the B.S. degree in computer engineering from Selcuk University, Konya, Turkey, the M.S. degree from the International Computer Institute, Ege University, Izmir, Turkey, and the Ph.D. degree in computer engineering from the Department of Computer Engineering, Ege University.

From 2011 to 2013, she was a Visiting Researcher at the Erik Jonsson School of Engineering and Computer Science, University of Texas at Dallas, Dallas, TX, USA. She is currently an Associate Professor at the Department of Computer Engineering, Ege University. She teaches courses in undergraduate and graduate computer engineering, advises graduate theses, and dissertations. She is also the coordinator of graduate thesis studies and a member of Semantic Web Technologies Research Group at the Department of Computer Engineering. Her research interests include access control mechanisms, privacy preserving techniques, knowledge engineering, and Semantic Web technologies.

...