Consent Through the Lens of Semantics: State of the Art Survey and Best Practices

Anelia Kurteva a,* , Tek Raj Chhetri a , Harshvardhan J. Pandit b , and Anna Fensel a

a Semantic Technology Institute (STI) Innsbruck, Department of Computer Science, University of Innsbruck, Innsbruck, Austria
b ADAPT Centre, School of Computer Science and Statistics Trinity College Dublin, Dublin, Ireland

Abstract. The acceptance of the GDPR legislation in 2018 started a new technological shift towards achieving transparency. GDPR put focus on the concept of informed consent applicable for data processing, which led to an increase of the responsibilities regarding data sharing for both end users and companies. This paper presents a literature survey of existing solutions that use semantic technology for implementing consent. The main focus is on ontologies, how they are used for consent representation and for consent management in combination with other technologies such as blockchain. We also focus on visualisation solutions aimed at improving individuals’ consent comprehension. Finally, based on the overviewed state of the art we propose best practices for consent implementation.


Keywords: Consent, GDPR, Semantic Web Technology, Ontology

1. Introduction

In the era of Big Data and the Internet of Things an unprecedented amount of data is being generated. According to the World Economic Forum1, the data generated by connected devices, social networking sites, including personal information, is a new asset in modern time [1]. However, when the data consists of sensitive and personally identifiable information, depending on the way it is used, the impact on the individual and the society at large could be both positive and negative [2]. The use of the data and the potential of harm (to fundamental rights such as privacy) is the principle behind laws such as the European General Data Protection Regulation (GDPR)2[3], which came into effect on 25th May 2018, superseding its predecessor - the Data Protection Directive (95/46/EC)3 and the national laws transposing it.

GDPR is designed to establish lawfulness, fairness and transparency regarding personal data processing. It is also designed for purpose and storage limitation, data minimisation, maintaining integrity, confidentiality and accountability. It applies to all individuals and organisations that collect and process information related to EU citizens, regardless of their location and data storage platform [4, 5]. The fines

*Corresponding author. E-mail: anelia.kurteva@sti2.at.
1https://www.weforum.org

0000-0000/21/$00.00 © 2021 –
for non-compliance with GDPR vary based on the severity of the law violations. According to GDPR the maximum fine is “up to 20 million euro, or 4% of the firm’s worldwide annual revenue from the preceding financial year, whichever amount is higher” (Article 83). In 2019 the National Commission on Informatics and Liberty (CNIL)\(^4\) fined Google with 50 million Euro for not complying with GDPR [6]. This action has set a warning and a strong message to all the technology companies about the consequences of not complying with GDPR. In order to avoid those fines, organisations must follow the six legal basis of GDPR, amongst which is consent implementation.

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” (Art. 4 (11)) and has introduced additional requirements for how consent should be collected. To be specific, consent must be:

- **Freely given.** Users have the right to consent or not based on the provided information. One should not be pressured to consent (Rec. 43).
- **Specific.** Consent should be requested about specific data (Art. 7).
- **Informed.** Users are presented with information about the data controller’s identity (Art. 7, Rec. 32).
- **Unambiguous.** Information should be provided in a “clear and plain” language (Rec. 42).
- **Could be withdrawn.** Users must be aware of their right to revoke consent. Further, the revocation option should be clearly stated and easily accessible. Revoking consent must be as easy as granting it from an end-user perspective (Art. 7 (3)), specifically w.r.t. the data to be processed, how it is to be used and the purpose of the processing.

The principle of consent is based on an individual’s agreement towards some specified action or intention. In practice, the use of consent as a legal basis for processing of personal data involves several relevant requirements and obligations which affect the interpretation of its validity. For example, informed consent requires provision of relevant information prior to consent. GDPR, being a pan-European regulation, redefined the use and practices surrounding consent by introducing a more stringent definition of consent along with additional requirements regarding the information to be provided and documented towards compliance.

In the context of GDPR, when consent is the legal basis, data processing can not begin before consent is obtained from the data subject. Any personal data processing without consent from the data subject (i.e. end-user) is liable for legal action defined by GDPR, highlighting its importance. Despite such importance of consent, to date, there is no single comprehensive collection of information describing requirements regarding consent across various relevant domains. Further, there is a lack of clarity regarding its implications in terms of legal compliance. This brings us to the questions such as how consent could be adopted in the future with the advancing use of technology without having to make many efforts, how the interpretation of privacy policies and visualisation of consent should be made and what the challenges associated with all these actions are. Therefore, there is a need for innovative consent implementation solutions that address the whole consent lifecycle (such as we have depicted in Figure 1) - from its representation, request, comprehension by users, decision-making by users (e.g. to give, to refuse, to withdraw consent) and its use (e.g. for compliance checking).

Semantic technologies, namely ontologies, have been gaining popularity in recent years due to their ability to specify and utilise relationships between entities and across domains and at large scales. On-

\(^4\)https://www.cnil.fr/en/cnils-missions
ologies allow a better knowledge discovery, interpretability, transparency and traceability of data [7–12]. Moreover, semantic web technologies are based on open and interoperable standards such as RDF (Resource Description Framework)\(^5\) for information representation, OWL (Web Ontology Language)\(^6\) for representation of ontological modeling and SPARQL\(^7\) for querying, and are extendable by design - making them suitable for application across use cases. In practice, due to the potential involvement of hundreds of organisations, consent implementation can develop into a complex ecosystem. Furthermore, the ability of semantic web technologies to model complex and dynamic ecosystems makes them suitable for consent implementation [13][14].

Otto et al. [15] present a survey of legal ontologies and approaches used in knowledge modeling. Their work helps to identify the role of various approaches for representation and legal compliance (e.g. deontic logic, symbolic logic, defeasible logic, temporal logic, access control) along with their strengths and weaknesses. The survey [15] informs how such ontologies can be used in different contexts such as modelling of the regulation itself or information for meeting compliance objectives of regulations. Further, Otto et al. [15] show that legal ontologies have been used in legal and regulatory compliance domains for quite some time.

The research by Rodrigues et al. [16] categorises legal ontologies along dimensions of (i) organisation and structuring of information, (ii) reasoning and problem solving, (iii) semantic indexing and search, (iv) semantic integration and interoperability and (v) understanding of a domain. The research in [16] shows that there are various approaches of legal domain and compliance that are addressed by ontologies and that they also assist in other knowledge and data driven processes.

Legal ontologies are also researched by Leone et al. [17]. The work in [17] investigates legal ontologies along several criteria with the aim of assisting “generic users” and legal experts in selecting a suitable ontology. The main domains of interest here are policies, licenses, tenders & procurements, privacy (including GDPR), and cross-domain (norms, legislations). The methodology in [17] includes the development and ontology engineering process, investigating use of ontological design patterns and reuse, and the relationship of modeling and concepts with legal norms and processes.

However, potential adopters of consent implementation solutions face the difficult question of identifying appropriate existing approaches, ontologies, the aspects of consent they model in terms of GDPR requirements, technical solutions, industry requirements and benefits and the peculiarities of design they utilise. In addition, investigations into whether these approaches can be used for different practical use cases, their scalability, efficiency and potential for adoption in changing requirements within the real-world remains a challenge. With this as the background and motivation, we present a survey comprising the state of the art for the implementation of consent as defined by the GDPR with the use of semantic technology.

The main contributions of our work can be summarised as follows:

- An overview of existing solutions for the semantic representation of consent and its management related to GDPR.
- An overview of graphical consent visualisation solutions aimed at raising one’s awareness regarding the implications of giving consent.
- An overview of relevant standardisation efforts.

\(^5\)https://www.w3.org/RDF/

\(^6\)https://www.w3.org/OWL/

\(^7\)https://www.w3.org/TR/rdf-sparql-query/
A set of best practices and recommendations for using semantic technology for consent representation, management and visualisation to end users.

The paper is organised as follows. Section 1 is an introduction to the topic, while Section 2 presents the followed methodology. Section 3 presents an overview of existing solutions in the fields of semantic models for consent, consent visualisation aimed at raising one’s awareness, consent management and current standardisation efforts. Based on the provided literature review, best practices for consent representation with semantic technology, management and visualisation are presented in Section 4. Conclusions are presented in Section 5.

2. Methodology

To create this paper, we followed a typical methodology for doing a survey, following the key principles of systematic reviews (PRISMA) [18]. We have selected the addressed areas, as well as the principles for the overviewed papers, projects and standardisation efforts. Given the motivation for this paper, the scope of work considered is defined as implementing consent (as defined by GDPR) with semantic technology. By implementing consent, we view the processes of consent modeling, consent management and consent visualisation.

Peer-reviewed publications were the primary source of knowledge regarding approaches, and were identified using the scholarly indexing services: Google Scholar [8], IEEE Xplore [9], ACM Digital Library [10], Scopus [11], and DBLP [12]. In addition to these, information was gathered through dissemination networks such as Twitter [13] and public mailing lists, standardisation-related websites, and information portals of the research funding agencies. Searches using keywords such as Consent Ontology, Informed Consent, Semantic Models for Consent, Consent Management Tools, Consent Visualisation, Consent Ethics, GDPR were used to identify relevant approaches in these sources. Authors and affiliations of identified publications were also used as keywords to find additional relevant resources. In cases where publications acknowledged funding or projects, an effort was made to identify its online website and access the list of publications. This provided information about the project’s aims and objectives, and its future goals and directions. The authors have also been participating themselves in the relevant European and nationally-funded projects, such as H2020 smashHit [14], FFG CampaNeo [15], FFG DALICC [16], and therefore had an insider view on the consent representation and modeling issues, and also found and analysed the information about the related projects on the websites of the funding agencies (European Commission, national funding agencies). Finally, relevant works at standardisation bodies have been overviewed.

In order to understand, analyse and categorise the approaches within the state of the art regarding its relation to consent, we introduce and use a model of ‘consent life-cycle’ (Figure 1). The consent life-cycle represents the different states and roles of information and semantics in processes associated
with consent. It consists of ‘Request’ as the state at which information must be provided for requesting informed consent, followed by ‘Comprehension’ where the individual must understand and interpret the provided information. ‘Decision’ consists of the individual (or agent) making a decision so as to give or refuse consent. Refusing consent requires it to be requested again, whereas giving consent permits its use to process data. ‘Consent Management’ is responsible (in addition to managing the request and collection of consent) to check the continued validity of consent to permit its use. Consent needs to be requested again if it is: withdrawn, expired, invalidated, revoked or it needs to be: modified, confirmed, or reaffirmed.

In each of these states, requirements related to internal organisational processes as well as legal compliance affect the information and processes involved, and therefore have an impact on the information and artefacts used to execute or implement them. For example, GDPR provides obligations regarding information to be provided to the individual (Art.13), which also affect information to be provided when requesting consent. For data controllers, this information must first be identified and then used to create a notice used in requesting consent. GDPR also provides obligations regarding the conditions and mechanisms for how consent should be requested which determine its validity as a legal basis (Art.7, Rec.32 and Rec.43). Therefore, the management of information related to consent is important for controllers as a matter of legal compliance. For individuals, the existence and presentation of this information affects its comprehension and therefore impacts the decision regarding consent for processing their personal data. A supervisory authority investigating compliance would want to ensure that the decision made by the individual is accurately represented and used to permit or prohibit the processing of personal data (Rec.42). Such investigations therefore involve information from all states in the life-cycle and can involve multiple industries. Thus, requirements derived from the consent life-cycle span across multiple domains and converge around the use of information. The use of semantics facilitates integration and interoperability of information across states and actors.

Our overview of existing work uses this as motivation to analyse and categorise approaches across fields in their relation to consent representation and management, and the potential for use of semantic technology. In particular, we consider (Section 3):

- Semantic models or ontologies for modeling information related to consent. Within this, we focus on the definition of consent as an ontological concept and other concepts and attributes that are associated with it.
• Approaches for management of information associated with consent, and its subsequent use to permit or prohibit processing.
• Approaches that aim to assist the individual regarding comprehension of information relevant to consent, with a particular focus on visualisation techniques.
• A discussion about relevant standardisation efforts.

Finally, analysing the state of the art from different angles relevant to consent representation, management and visualisation, we identify the current challenges and gaps, as well as the best practice recommendations for the consent modeling, management and visualisation, that are of benefit to the research, developer and practitioner communities. When doing so, we additionally take into account ethical and sociological aspects regarding practices surrounding consent, and its impact on individuals.

3. Overview of Related Work

This section provides an overview of related work in the areas of consent modelling, graphical visualisation of consent to end users, consent management and current standardisation efforts. We view consent representation from a semantic perspective and present semantic models for consent, namely ontologies. Next, we provide an overview of work on graphical consent visualisation to end users aimed at raising one’s awareness regarding the implications of giving consent. Further, various existing and developing solutions for consent management based on semantic technology are presented. Finally, a short summary of current standards for consent is presented as well.

3.1. Semantic Models for Consent

Ontologies are some of the most essential semantic web technologies used for representing concepts and the relationships between them in both human-readable and machine-readable formats. Some of the reasons for using ontologies are: to share common understanding of the structure of information among people or software agents, to enable reuse of domain knowledge, to make domain assumptions explicit, to separate domain knowledge from operational knowledge, and to analyse domain knowledge. In the case of consent, an ontology provides a formal conceptualisation that is interpretable by the different entities involved in the data sharing process. We view a semantic model as a consent ontology, if as a minimum, the concepts of consent and its purpose are modelled.

This section provides an overview of consent ontologies by stating (i) the purpose of the ontology, (ii) language used for specification, (iii) how consent is modelled, and (iv) level of detail when modeling personal data for consent (e.g. presence of abstract or specific instances, granularity of concepts, specific taxonomies or instances, domain-specific or use case specific). Further, we used a set of competency questions (Table 1) for evaluating to what extent each ontology is capable of representing information regarding informed user consent. The competency questions were derived by following GDPR requirements for informed consent and already existing sets of competency questions such as the one of G Consent\[17\]. The ontologies reviewed in this section are CDMM\[18\], G Consent\[17\], PrOn  to[19], LloPY[20], BPR4GDPR\[23\], SPL and SPL Log [21], ColPri [22] and DPV\[27\].

\[17\]http://openscience.adaptcentre.ie/ontologies/GConsent/docs/ontology
3.1.1. Consent and Data Management Model (CDMM)

The CDMM\(^{18}\) ontology by Fatema et al. [23] utilises a consent ontology written in OWL\(^{6}\). The ontology represents a generic model for consent, permissions and prohibitions according to the GDPR and further reuses the PROV-O\(^{19}\) ontology to express provenance information from different systems [23]. CDMM allows to represent the format in which consent was retrieved such as app based, audio, online form, etc. Keeping track of changes in the state of data, consent and operations is made possible by defining the classes for time, use and action. The ontology models both personal data, such as health data, and non-personal data i.e. any data that is not sensitive according to the given consent. Further, CDMM provides classes for different data formats such as video, audio, picture, text and defines three types of processing (examine, modify and read). CDMM can be used for consent management (e.g. collecting consent, maintaining records of consent).

3.1.2. GConsent

GConsent\(^{17}\), an ontology written in OWL2\(^{20}\), models information about consent based on requirements of GDPR compliance [24]. It represents consent as an artefact that can have states indicating its lifecycle - such as requested, given, refused, or withdrawn. The relevant information regarding purpose, personal data categories, processing, and parties involved are associated with a central concept representing "consent". Novel aspects of this ontology involve modeling of the context in which consent was requested or given, such as location and medium. The ontology also provides representation of delegation regarding consent, and provides examples of its application in several use-cases. For example, GConsent can be used when modelling information (e.g consent) related to GDPR compliance.

3.1.3. Privacy Ontology (PrOnto)

The PrOnto ontology [19], written in OWL\(^{6}\), is used for modelling GDPR concepts such as privacy agents, data types, types of processing operations, rights and obligations. Consent is viewed as one of the legal bases used to justify a processing activity. PrOnto models the concepts for purpose, personal data (e.g. health, genetic, ethnic, sexual data), and non-personal data (e.g. anonymous data) in its data model and associates them with a legal basis. The structure of the ontology is based on five modules: (i) documents and data, (ii) actors and roles, (iii) processes and workflow, (iv) legal rules and deontic formula, (v) purposes and legal bases. The ontology provides a significant number of concepts (for combining different ontologies and design patterns) for modelling GDPR-related concepts, but also strives to go beyond the GDPR requirements so that it could be applied in any legal scenario. For example, the ontology can be used for compliance checking during the whole lifecycle of the personal data. [19].

3.1.4. Legal Complaint Ontology to Preserve Privacy for the Internet of Things (LloPY)

The LloPY [25] ontology, developed with OWL and aimed to be used in the Internet of Things (IoT), follows the NIST (National Institute of Standards and Technology Interagency Report)\(^{21}\) privacy definition. Consent is viewed from a privacy perspective and is represented as a privacy attribute. The privacy attributes are derived based on GDPR and NISTR [26]. LloPy models the purpose for consent, retention, disclosure, operation, condition, etc. The ontology is utilised by the IoT Resource Management Module of the system presented in [25], which performs data anonymisation, noise addition, etc. In addition to modelling, consent for privacy preservation in smart devices, LloPY reuses the Semantic Sensor

\(^{18}\)http://purl.org/adaptcentre/openscience/ontologies/consent

\(^{19}\)https://www.w3.org/TR/prov-o/

\(^{20}\)https://www.w3.org/TR/owl2-overview/

Network ontology (SSN)\(^{[22]}\), which provides more detailed privacy properties for sensors and their observations. The ontology can be used when one needs to model consent for sensor data sharing in the IoT, for example, in smart cities.

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
<th>Relevant Concept(s)</th>
<th>Relevant GDPR Clause(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Who collects the data?</td>
<td>Data Controller, Data Processor</td>
<td>Art. 4 (7), Art. 6, Art. 28</td>
</tr>
<tr>
<td>2</td>
<td>For what purpose?</td>
<td>Purpose</td>
<td>Art. 4 (4), Art. 6 (1a, 1f, 4), Art. 7 (32)</td>
</tr>
<tr>
<td>3</td>
<td>How to withdraw consent?</td>
<td>Consent Withdrawal</td>
<td>Art. 17, Rec. 63, Rec. 66</td>
</tr>
<tr>
<td>4</td>
<td>How long does consent last for?</td>
<td>Consent Duration/Validity/Expiry</td>
<td>Rec. 32, Rec. 42</td>
</tr>
<tr>
<td>5</td>
<td>When was consent given/revoked?</td>
<td>Consent Duration/Revocation</td>
<td>Art. 17, Art 19</td>
</tr>
<tr>
<td>6</td>
<td>What personal data is collected?</td>
<td>Personal Data Categories</td>
<td>Art. 4 (1), Art. 9</td>
</tr>
<tr>
<td>7</td>
<td>How is the personal data being used?</td>
<td>Processing</td>
<td>Art. 4 (2)</td>
</tr>
<tr>
<td>8</td>
<td>How is personal data collected?</td>
<td>Data Collection</td>
<td>Art. 12, Art. 13, Art. 14, Rec. 39, Rec. 58, Rec. 62, Rec. 73</td>
</tr>
<tr>
<td>9</td>
<td>With whom is personal data shared?</td>
<td>Recipient, Data Sharing</td>
<td>Art. 4 (7), Art. 6, Art. 28</td>
</tr>
<tr>
<td>10</td>
<td>Who is responsible for the personal data?</td>
<td>Data Controller</td>
<td>Art. 24, Rec. 74, Rec. 79</td>
</tr>
<tr>
<td>11</td>
<td>Where is personal data stored?</td>
<td>Data Storage</td>
<td>Art. 5</td>
</tr>
<tr>
<td>12</td>
<td>Who is the Data Controller?</td>
<td>Data Controller</td>
<td>Art. 4 (7), Art. 28</td>
</tr>
<tr>
<td>13</td>
<td>How to contact the Data Controller?</td>
<td>Data Controller, Contact Inform</td>
<td>Art. 4 (7), Art 14, Art. 28</td>
</tr>
<tr>
<td>14</td>
<td>What are the responsibilities of the Data Controller?</td>
<td>Data Controller, Responsibilities, Obligations</td>
<td>Art. 4 (7), Art 14, Art. 28, Art. 37</td>
</tr>
<tr>
<td>15</td>
<td>Who is the Data Subject?</td>
<td>Data Subject</td>
<td>Art. 4 (1)</td>
</tr>
<tr>
<td>16</td>
<td>Whom to contact?</td>
<td>Contact Information, Third Party</td>
<td>Art. 12, Art. 13, Art. 14</td>
</tr>
</tbody>
</table>

### 3.1.5. Business Process Re-engineering and Functional Toolkit for GDPR Compliance (BPR4GDPR)

The compliance ontology developed as deliverable D3.1\(^{[23]}\) of the BPR4GDPR\(^{[24]}\) project aims to provide the fundamental entities, concepts and relationships that are needed for achieving compliance. The ontology was built based on project work done in the legal and technical fields and has a hierarchical data type structure, which allows for the detailed organisation of entities and interrelations. Amongst the core concepts in the ontology are roles, event types, context types and state types. Further, the ontology models the concept of a purpose, which is a GDPR requirement for informed user consent. Having

---

\(^{[22]}\)https://www.w3.org/TR/vocab-ssn/


\(^{[24]}\)https://www.bpr4gdpr.eu
Table 2
Overview of Existing Semantic Models for Consent

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Year of latest update</th>
<th>Availability</th>
<th>Scope</th>
<th>How is consent modelled/viewed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDMM</td>
<td>2017</td>
<td>Open-access</td>
<td>Data provenance</td>
<td>Consent is viewed as an entity within a privacy policy.</td>
</tr>
<tr>
<td>GConsent</td>
<td>2018</td>
<td>Open-access</td>
<td>GDPR compliance</td>
<td>Consent is modelled as an artefact, which has states (given, not given, refused, withdrawn).</td>
</tr>
<tr>
<td>PrOnto</td>
<td>2018</td>
<td>Private</td>
<td>GDPR obligations and require-</td>
<td>Consent is viewed as one of the legal bases used to justify a processing activity.</td>
</tr>
<tr>
<td>LloPy</td>
<td>2018</td>
<td>Private</td>
<td>Privacy and security</td>
<td>Consent is modeled from a privacy perspective as an attribute.</td>
</tr>
<tr>
<td>BPR4GDPR</td>
<td>2019</td>
<td>Private</td>
<td>GDPR compliance</td>
<td>Consent is modeled as an event type (provided, revoked, refused).</td>
</tr>
<tr>
<td>SPL and SPLog</td>
<td>2019</td>
<td>Open-access</td>
<td>GDPR compliance</td>
<td>Consent is modelled as a policy and is used for compliance checking.</td>
</tr>
<tr>
<td>ColPri</td>
<td>2020</td>
<td>Private</td>
<td>Privacy policies in the IoT</td>
<td>Consent is modelled as a privacy policy and has two states (given and ungiven).</td>
</tr>
<tr>
<td>DPV</td>
<td>2021</td>
<td>Open-access</td>
<td>Privacy and legal compliance</td>
<td>Consent and its attributes (e.g. expiry time) are modelled as privacy policies for cases such as personal data handling and compliance checking.</td>
</tr>
</tbody>
</table>

such diversity of data types allows to define consent in detail and a precise compliance check to be performed. The ontology can be used for modelling consent as an access control for compliance checking. Full specification of the Compliance Ontology is available in Deliverable D3.1\(^{23}\) of the BPR4GDPR project.

3.1.6. SPECIAL’s Usage Policy Language (SPL)

The SPECIAL’s Usage Policy Language (SPL)\(^{21}\), developed for the SPECIAL-K compliance platform, is a language for modeling usage policies. SPL encodes the usage policies in OWL2. SPL models data processing, the purpose for processing, description of the operations and the involved entities. A detailed description of the SPL ontology can be found in deliverable D2.1\(^{27}\). The SPL’s scope is limited to capturing the permissive nature of given consent in order to compare it with its processing logs to determine (and evaluate) compliance according to the given consent. However, the vocabulary also models purpose, processing, recipients, temporal duration, etc. The main aim of the language is to model data subject’s consent and relevant data usage policies in a machine-readable formal way, and to define permissions based on the given consent thus allowing compliance checking and policy verification\(^{21}\).

The SPLog\(^{25}\) vocabulary builds upon the existing SPL by reusing existing vocabularies for data provenance such as PROV\(^{19}\) and represents consent states such as revocation and assertion as types of "PolicyEntry". The class "ConsentAssertion" defines the consent received by the data subject, while "ConsentRevocation" models the action of consent revocation. These two classes, being subclasses of "PolicyEntry", which is also a subclass of "LogEntry" allow for the direct linking of consent to the data subject and vice versa. Both vocabularies can be used for modelling consent as system logs in privacy policies in order to restrict data usage and processing\(^{21}\).

\(^{23}\)https://ai.wu.ac.at/policies/policylog/
A. Kurteva et al. / Consent Through the Lens of Semantics: State of the Art Survey and Best Practices

Table 3
Overview of Existing Semantic Models for Consent: Classes and Properties representing Consent

<table>
<thead>
<tr>
<th>Ontology</th>
<th>Classes</th>
<th>Object Properties</th>
<th>Relevant Consent Lifecycle Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDMM</td>
<td>Consent, ConsentFormat, ConsentingParty, ConsentObligation</td>
<td>consent_given_at, consent_given_by, consent_given_for, data_has_format</td>
<td>Request, Comprehension, Decision, Use</td>
</tr>
<tr>
<td>GConsent</td>
<td>Consent, Data Subject, Personal Data, Processing, Purpose, Status, Expired, Explicitly Given, Given by Delegation, Implicitly Given, Invalid, Not Given, Refused, Requested, Unknown, Withdrawn</td>
<td>hasStatus, hasConsent, isActionForPurpose, isContextForConsent, isPersonalDataForConsent, isPreviousConsentFor, isPurposeForConsent, isConsentGivenFor, isConsentGivenBy, isConsentGivenFor, data_has_format</td>
<td>Request, Comprehension, Decision, Use</td>
</tr>
<tr>
<td>BPR-GDPR</td>
<td>ConsentProvided, ConsentRevoked, ConsentDenied, DataProcessor, DataSubject, DataController, DataProtectionAuthority, DataProtectionOfficer</td>
<td>hasData, hasProcessing, hasPurpose, hasRecipient, hasController, hasProvisionBy, hasProvisionTime, hasWithdrawalTime, hasWithdrawalByJustification, hasWithdrawalMethod, hasWithdrawalTime, isExplicit</td>
<td>Request, Decision, Comprehension, Use</td>
</tr>
<tr>
<td>SPL and SPLog</td>
<td>LogEntry, PolicyEntry, ConsentAssertion, ConsentRevocation</td>
<td>hasData, hasProcessing, hasPurpose, hasRecipient, log_controller, log_revoke, log_revocation, at_time, is_provided_to_controller</td>
<td>Request, Decision, Comprehension, Use</td>
</tr>
<tr>
<td>DPV</td>
<td>Consent, Purposes, LegalBasis, DataSubject, DataController, Right</td>
<td>hasConsentNotice, hasExpiry, hasExpiryCondition, hasExpiryTime, hasProvisionBy, hasProvisionByJustification, hasProvisionMethod, hasProvisionTime, hasWithdrawalTime, hasWithdrawalByJustification, hasWithdrawalMethod, hasWithdrawalTime, isExplicit</td>
<td>Request, Decision, Comprehension, Use</td>
</tr>
</tbody>
</table>

3.1.7. Collaborative Privacy Knowledge Management Ontology for the Internet of Things (ColPri)

The ColPri ontology [22], developed with OWL\(^6\) and using the SKOS\(^{26}\) vocabulary, aims to provide a collaborative IoT knowledge base which enables one to configure privacy policies. Consent is viewed from a privacy perspective and is modeled as a privacy attribute with two states: *given* and *ungiven*. The purpose of consent is defined as either *Advertising* or *ApplicationFunctioning*. Further, the ontology allows one to specify if information disclosure to entities such as developers and third parties is allowed. Regarding personal data, ColPri follows SKOS and models different data categories such as personal, pseudonymous, and anonymous data. Personal data could be further specified as sensitive (e.g. criminal, health, habit and identity) and nonsensitive. ColPri differs from other ontologies by using both OWL and SKOS thus allowing flexible data categorisation and privacy policy handling based on user consent. The ontology can be used for modeling data privacy preferences in smart cities, specifically in smart homes.

3.1.8. Data Privacy Vocabulary (DPV)

The Data Privacy Vocabulary (DPV)\(^{27}\), is an outcome and deliverable of the W3C Data Privacy Vocabulary and Controls Community Group (DPVCV)\(^{28}\). The DPVCV was formed as an activity of the SPECIAL project, and represents a broad consensus amongst experts from the domains of data protection, privacy, legal compliance, and semantic web. DPV provides a vocabulary of concepts based primarily on GDPR, along with hierarchical top-down taxonomies for specifying purposes, processing categories, personal data categories, technical and organisational measures, and GDPR’s legal basis (as an extension called DPV-GDPR). The representation of consent in DPV is through the concept *Consent* along with properties enabling representing notice, expiry, provision, withdrawal, and whether it is ex-

---

\(^{26}\)https://www.w3.org/2004/02/skos/

\(^{27}\)https://w3.org/ns/dpv

\(^{28}\)https://www.w3.org/community/dpvcg/
The association of purposes, processing, personal data categories and other relevant information is represented through the `PersonalDataHandling` class which associates consent as the legal basis used for a particular instance of processing. The modeling of consent within DPV is based on the requirements of GDPR for recording and documenting given consent and the `Consent Receipt` specification. DPV can be used for representing responsibilities and obligations in privacy policies and to "support automated checking of legal compliances of data handling ex ante (prior to processing), or ex post (i.e. check compliance after processing)".  

### 3.1.9. Summary

A summary of the ontologies that were discussed in this section, their scope and the way each one models consent is presented in Table 2. The specific classes and object properties used for modelling consent, for each ontology (based on resources available online) from Table 2 are presented in Table 3.

Table 4 presents the evaluation of the ontologies from Section 3, with the competency questions from Table 1. A "check sign" (✓) is used if the ontology is able to answer the question (i.e. the concept is present in the ontology), and an empty space is used where concepts were not found, while acknowledging they could be added later e.g. through an update. The findings show that the existing ontologies are quite diverse based on their scopes and when it comes to their abilities to model consent.

GConsent, SPL [28] and BPR4GDPR are aimed at modeling consent while taking into account GDPR requirements. DPV also models consent (from privacy perspective), but the main focus if on GDPR as a whole. PrOnto [19], ColPri [22] and LloPY [25] are developed from a privacy perspective and view consent as an attribute that helps preserve data privacy. Similarly, CDMM models consent as an entity within a privacy policy and further allows for the capturing of data provenance. From a technical standpoint, the OWL standard is followed, with an exception of the ColPri ontology which further utilises the SKOS organisation system. Regarding the ability to represent informed user consent, the ontologies reviewed in this section are still somewhat generic, have a specific scope (Table 2) and achieving such level of detail while being compliant with GDPR requires combining several ontologies. By far, GConsent, PrOnto and BPR4GDPR have the potential to be both GDPR compliant and to represent informed user consent in detail. In conclusion, various ontologies for consent have been developed in the past, however, common limitations are present.
3.2. Consent Visualisation

When talking about consent and its representation with semantic technology, one should also consider how it is visualised (e.g. via a user interface (UI) or graphically) to the end users in an informative way as no process can start without one’s consent. However, having users’ informed consent does not mean that the user understands the consequences of his or her action. The desire for convenience, fast and easy interactions may make one disregard important information regarding consent and simply agree to anything that is required without being aware of the consequences. Bechmann [29] defines this as a "culture of blind consent". The issue is also addressed by Joergensen et al. [30] who examined the user’s understanding of privacy policies, data control and the importance of social media as a whole. The results showed that the need to be accepted is enough to influence users to consent. Users had a general common sense of what types of information should and should not be shared online but they lacked knowledge regarding data sharing on a company level and the related privacy risks. The study validated Bechmann’s point [29] that users lack knowledge about what it means to consent and that they are more concerned with how they would be perceived by others. Human Computer Interaction (HCI) [31] is a broad field by itself thus we limit the scope of this section to research and projects that focus specifically on visualising informed user consent (via a UI) to raise one’s awareness. An overview of the following UIs is presented below: Data Track [32], The Privacy Dashboard [33], CoRe [34], CURE [35].

3.2.1. Data Track

Angulo et al. [32] developed a tool for visualising data disclosures called Data Track (Figure 2). The tool’s development was initially part of the European PRIME[29] and PrimeLife[30] projects and then continued as part of the A4Cloud[31] project. The motivation for the tool is to enable transparency and raise awareness regarding what is happening to one’s data. Data Track’s main goals are to allow users (i) to monitor how their data is being used by different online services and (ii) to exercise their rights. Monitoring of the data flow is achieved by providing users with a graphical visualisation, which the authors refer to as “trace view”. The main concept of the trace view is that the user is at the center of everything thus making one feel as if the interface focuses on them. The interface itself is divided in two panels. The bottom panel allows one to view information provided to each service, while the top one displays the information currently being shared. Further, upon selecting a specific service a user is presented with a new window displaying a more detailed overview of what data is being shared and is given the possibility to edit permissions. Users deemed the interface as useful as it helped them become more aware of what is happening to their data. However, the evaluation showed that even users, who were knowledgeable about the web, lacked understanding about how their data is collected, shared and used.

3.2.2. The GDPR-compliant and Usable Privacy Dashboard

Raschke et al. [33] develop a privacy dashboard that enables users to execute their rights according to GDPR. The implementation of the user interface follows Nielsen’s Usability Engineering Lifecycle [36]. The authors start by analysing the user’s and the tasks they need to complete and then develop several parallel versions of the privacy dashboard. The prototype (Figure 3), namely a single page that consists of three main building blocks (general functionalities, data overview and general information),
Fig. 2. The Data Track Tool by Angulo et al. [32]

was developed with JavaScript and React. The general functionalities plane allows the user to review given consent, request information about involved entities, view privacy policies, etc., while the data overview plane visualises the data flows with the help of an interactive graph, which is implemented with the vis.js library. The general information section, located on the right-side of the dashboard, provides details about third-parties such as name and address. The privacy dashboard has proved to be useful as it made users more aware about their rights. The authors suggest that future improvements of the design to minimise information overload are needed [33].

3.2.3. The CoRe and CURE User Interfaces

Drozd and Kirrane [34][35] address consent and the challenge of its representation to end-users by developing the CoRe UI [34] (Figure 4) and its third iteration called CURE [35] (Figure 5). The CoRe UI is based on GDPR requirements and aims to minimise the issue of information overload that is present in existing solutions. As discussed there, most of the existing work is focused on developing GDPR privacy policies and not on the representation of consent and its visualisation to the end user, thus a new methodology for achieving this is presented. The methodology is based on the Action Research (AR), which requires a problem to be defined first. Following a sample use case, several UI prototypes were devel-
Table 5
Graphical Consent visualisation via a UI

<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>What is visualised</th>
<th>How is it visualised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Track</td>
<td>2015</td>
<td>Personal data processing, user rights.</td>
<td>Personal data and its processing is visualised with a tracing graph on a UI.</td>
</tr>
<tr>
<td>The Privacy Dashboard</td>
<td>2018</td>
<td>Consent, data privacy rights, processing.</td>
<td>A UI enables the chronological and interactive graphical representation of data processing.</td>
</tr>
<tr>
<td>CoRe and CURE</td>
<td>2019</td>
<td>Consent, purpose, data, storage, processing, sharing.</td>
<td>Consent requests are visualised on a UI with the help of interactive graphs.</td>
</tr>
</tbody>
</table>

The CoRe UI by Drozd and Kirrane [34]

![CoRe UI](image)

Fig. 4. The CoRe UI by Drozd and Kirrane [34]

opened with Angular\(^{32}\) and D3.js\(^{33}\) and then tested with users both remote and onsite. Regarding consent representation, the “all or nothing” approach is put aside and users are given full flexibility to customise their consent. The UI enables users to explore possible consent paths via a hierarchical visualisation done with D3.js and to select a specific one they wish to follow. Further, understandability is addressed by avoiding the commonly used legal jargon and instead focusing on simple sentence structure.

What differentiates the CURE UI [35] (Figure 5) from other interfaces and consent forms is that it focuses on mobile device interaction and personalisation. Users have full control over their consent specification and data. In comparison to CoRe [34], that is based on the AR methodology, CURE follows the Design Science Research (DSR) paradigm, which is usually used for improving existing software [35]. The front-end was developed with Angular and D3.js, while Java\(^{34}\) and PostgreSQL\(^{35}\) were used on the back-end. Similarly to CoRe, the main objectives of the CURE UI are customisation, understandability and revocation. Customisation is achieved by allowing users to select what information they want to receive/share (e.g. health data) and for which purposes. In addition to using, as described, “simple” phrases, the UI provides users with feedback on demand upon each interaction in order to minimise the data overload and help understandability. Further, as in CoRe, a graphical representation of the consent process is provided. Consent revocation is done either by sliding the pointer up or by deselecting some of the options.

---

\(^{32}\)https://angular.io  
\(^{33}\)https://d3js.org  
\(^{34}\)https://www.java.com  
\(^{35}\)https://www.postgresql.org
3.2.4. Summary

The work on the CoRe [34], CURE [35], The Privacy Dashboard [33] and the Data Track [32] UIs (see Table 5) show that visualisation helps to raise one’s awareness about consent and the implications that follow. In addition, visualisation of the data helps achieve transparency, which is key for making well-informed decisions such as giving consent.

3.3. Consent Management

Having modeled consent semantically and visualised it graphically to the end user, one should next consider how to manage it. However, one can also consider or wish to manage consent without visualising it. Consent management could be viewed from both individual and system perspective, however, both are interlinked. While users must be able to perform actions such as giving and withdrawing consent at any time, the system must be able to handle them. Consent management, as defined by Pallas and Ulbricht [37], is a collection of processes that “allow or integrate queries upon multiple and autonomous data sources, taking into account data subjects’ individually given, purpose- and utilizer-specific, and dynamically adjustable consent”. Consent management, in most cases, refers to the controller managing the state or processes associated with consent in terms of whether it has been requested and obtained for the intended purposes and processing of personal data associated with it. It also refers to the use of (given) consent as permissions or access control to control the processes based on it. From a legal compliance perspective, consent management also refers to evaluating and maintaining the validity of consent and its associated processes based on obligations derived from law. The individual’s perspective involves tracking what consent was given, its withdrawal for the same set of information. Evidently, the processes should be adequately designed. Such a consent management system should particularly take into account the current policies and laws that need to be followed [38]. In the context of GDPR, consent management must comply with the obligations for personal data processing that are defined in GDPR’s Chapter 2 (Art. 5-11). For example, consent management operating within the EU or dealing with EU citizens must follow GDPR directives such as “Lawfulness of processing”, “Conditions for consent”, etc. as described in Art. 6, Art. 7 respectively. This section describes technological solutions for consent management that assist in the storage, use, evaluation, and documentation of consent based on requirements of GDPR compliance. We begin by providing an overview of each solution by specifying its scope, main goals and the motivation behind it. Next, we provide information about how consent management is achieved, followed by possible real-world applications.
3.3.1. EnCoRe

EnCoRe\textsuperscript{36} is a collaborative project between researchers in the UK that aims to develop a mechanism for consent revocation that could be successfully adopted by any business, and for raising awareness regarding one’s rights over their personal data. Regarding the architecture of the solution, the "Personal Consent and Revocation Assistant" provides users with the opportunity to give consent or revoke consent via a user interface, which also keeps record of one’s actions. Upon giving consent, the user data is sent to a virtual instance of a database called "Virtual Data Registry" and is further managed with the help of the Data Viewer and Manager component. Prohibitions, obligations and permissions are defined by the Privacy-aware Policy Enforcement, which together with the Disclosure and Notification Manager keep track of changes in the data flow. Changes in the state of the consent are recorded by the "Audit" component. The "Trust Authority" deals with compliance checks and certification of digital certificates, while the "Risk Assurance" component, which could be used offline as well, provides insights about security and privacy risks and suggestions on how to avoid them.

3.3.2. ADvoCate

ADvoCATE \textsuperscript{39} is a consent management platform based on blockchain technology, with the goal to provide information about data, detect violations of privacy policies and manage the data processing in an Internet of Things (IoT) ecosystem \textsuperscript{39}. The platform is used as a medium between the end-user and the industry and consists of (i) a consent management component, (ii) a consent notary component, and (iii) an intelligence component. Consent representation, updates and withdraws are managed by the consent management component with the data protection ontology by Bartollini et al. \textsuperscript{40} according to GDPR requirements. The consent notary component ensures compliance and consent validity by using reasoning, supported by Fuzzy Cognitive Maps (FCM), over the Ethereum blockchain, which manages the integrity and the versioning of consent, while the intelligence component identifies conflict in personal data sharing policies with the help of Fuzzy Cognitive Maps (FCM) \textsuperscript{41}, the Intelligent Policies Analysis Mechanism (IPAM) and the Intelligent Recommendation Mechanisms \textsuperscript{39}. The final solution is a framework that is able to record, validate and store user consent by combining semantic technologies, namely ontologies, and blockchain. The primary use of blockchain in the project is (i) for smart contracts, which are signed digitally using private key and (ii) for managing hashes. The mapping of data can be performed by using the unique id provided for each IoT device, which has been registered in the ADvoCate platform. The authors conclude that a more detailed ontology for consent and improvements of the intelligence component will be needed in the future.

3.3.3. SPECIAL-K

The SPECIAL-K is a framework developed under SPECIAL\textsuperscript{37} (Scalable Policy-aware Linked Data Architecture For Privacy, Transparency and Compliance) EU H2020 project for automatic compliance verification based on usage control policies for data processing and sharing. The motivation comes from the lack of consent management solutions that successfully execute its withdrawal. The main goal of the project is thus to have a framework that monitors consent and enables actions such as withdrawal to be immediately executed even after years of data sharing, while being compliant with current laws \textsuperscript{21}. The framework in \textsuperscript{21} consists of three primary SPECIAL components: (i) Consent Management Component, (ii) Transparency and Compliance Component, and (iii) Compliance Component. The Consent Management Component is responsible for obtaining consent from the data subject and representing

\textsuperscript{36}https://www.hpl.hp.com/breweb/encoreproject/index.html

\textsuperscript{37}https://www.specialprivacy.eu/
using SPECIAL usage policy vocabulary [21]. The Transparency and Compliance Component is responsible for presenting data processing and sharing events to the user following SPLog vocabulary (Section 3.1.6). The Compliance Component focuses is used to verify the compliance of data processing and sharing with usage control policies.

The implementation uses SPL[38], which is encoded using web ontology language (OWL2) to represent the policies, MongoDB[39] to store data about consent, embedded HermiT[40] reasoner to determine the compliance based on usage control policies, Elasticsearch[41] for browsing logs serialised using JSON-LD and Apache Kafka[42] to carry out processing of application logs and to save the result of reasoning in new Kafka topic.

3.3.4. GDPR Compliance Privacy Framework by Davari et al.

Davari et al. [20] present a GDPR privacy protection framework for an access control system that utilises XACML (an OASIS standard for expressing policies). The main aim of the research is to provide a solution that supports data privacy protection based on GDPR. The presented compliance validation model uses the PROV-O[19] ontology for semantically modelling consent according to GDPR. The consent model itself is built by extracting all GDPR relevant rules. The management of the consent and the personal data is done by utilizing the blockchain framework Hyper-ledger Fabric[43]. For imposing consent on all entities involved in the data sharing process, the authors use cryptography technology. Each party involved, such as the data subject, data processor, the data controller, is assigned an asymmetric key pair, and it is used as an identity mechanism. However, in addition to blockchain, MongoDB[39] is used for storing data. The main reason, as explained by Davari et al. is that blockchain is immutable thus data cannot be deleted once stored. Although this supports traceability and transparency, it is in collision with the user’s right to “erasure” given by GDPR.

3.3.5. CampaNeo

CampaNeo[44], a German-Austrian collaboration project with duration of three years (2019-2022) that aims to develop a platform for sensor data sharing between multiple entities. The platform’s main goal is to provide the industry with an outlet for publishing data requests for user’s vehicle sensor data in the form of campaigns. CampaNeo utilises machine learning for detection of driving behaviour, finding driver’s efficiency scores, predicting car accidents, traffic regions etc. and knowledge graphs for the campaign data modelling. The CampaNeo ontology defines the concepts of campaign, data, processing, third-party entities, users and consent. Knowledge graphs are used for achieving process transparency and data traceability by recording consent and its provenance. Further, a UI that focuses on consent visualisation with the help of forms is being currently developed (as of 2020). The UI aims to present users with information about consent such as its purpose, data regarding it, the organisation making the request, thus achieving GDPR compliance.

3.3.6. Blockchain-based Consent Model by Jaiman et al.

Jaiman et al. [42] present a dynamic GDPR consent model for health data sharing in a distributed environment, that utilises blockchain. The main motivation for their work is improving accountabil-

---

[38]https://www.w3.org/community/dpvcg/wiki/SPECIAL_usage-policy/
[39]https://www.mongodb.com/
[41]https://www.elastic.co/elasticsearch/
[42]https://kafka.apache.org/
[44]https://projekte.ffg.at/projekt/3314668
ity in health data sharing, which has proven to be a challenge due to the large volumes of data constantly being collected by consumer wearables. The developed blockchain-based consent model reuses the Data Use Ontology (DUO)\(^{45}\), which allows describing data use conditions for research data in the health/clinical/biomedical domain. Further, Jaiman et al. [42] reuse the Automatable Discovery and Access Matric (ADA-M) [43] ontology for classifying data use conditions and permissions. The consent statement itself is modelled with DUO then saved as a smart contract and added to the existing blockchain. Upon a data request from a third party, the ADA-M ontology is used for finding matching contracts. Once a match between the user consent statement and the data request is found access is granted to the requestor. When it comes to specific technology, the Solidity language for smart contracts and the LUCE platform for data sharing, which builds upon the Ethereum\(^{46}\) blockchain, were used [42].

3.3.7. Automated GDPR Compliance using Policy Integrated Blockchain by Mahindrakar et al.

Mahindrakar et al. [44] present a blockchain-based approach to facilitate GDPR compliance for real-time automated data transfer operations between consumers and providers. The main aim of their work is to ensure valid data transfer operations while maintaining GDPR compliance. The presented work uses both semantic technology and blockchain. Two ontologies are used, namely a GDPR ontology built by the authors and the privacy policy ontology by Joshi et al. [45], which represents consent from a privacy perspective. Management of consent, namely its validation, is done by querying the privacy policy ontology by Joshi et al. [45] using SPARQL\(^7\) and based on the result, further processing (e.g. data transfer) is allowed or not. The developed GDPR ontology by Mahindraker, itself, holds the information about GDPR articles. The relevant articles between consumers and providers are queried using SPARQL to create a GDPR knowledge graph, which is then used for reasoning with smart contracts. Regarding the implementation, the solution uses Natural Language Processing (NLP) techniques, the private blockchain network Ganache-CLI\(^{47}\) for Ethereum and encryption mechanisms (i.e. The Advanced Encryption Standard algorithm). Similarly to Davari et al. [20], the authors address the issue of the immutability of blockchain and how it affects GDPR compliance. To overcome this, data is saved in an external encrypted file, which is stored in a relational database. All the involving parties are registered on the blockchain network and are assigned a unique account number and a private key. By decrypting using the public key, the data owner is able to use the transaction hash stored in an encrypted file to access the transaction details.

3.3.8. smashHit

smashHit\(^{48}\) is an ongoing Horizon 2020 project that ends in December 2022 with the primary objective of creating a secure and trustworthy data sharing platform with focus on consent management in a distributed environment such as the automotive industry, insurance and smart cities. smashHit proposes to use semantic models of consent, such as ontologies and knowledge graphs and legal rules for consent management. The vision of smashHit is to overcome obstacles in the rapidly growing data economy, which is characterised by heterogeneous technical designs and proprietary implementations, locking business opportunities due to the inconsistent consent and legal rules among different data-sharing platforms actors and operators.

\(^{45}\)http://www.obofoundry.org/ontology/duo.html
\(^{46}\)https://ethereum.org/en/
\(^{48}\)https://www.smashhit.eu
Table 6

<table>
<thead>
<tr>
<th>Project/research work</th>
<th>Duration</th>
<th>Use Case</th>
<th>How is technology used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>EnCoRe</td>
<td>2008-2011</td>
<td>An end-user discloses personal data along with consent/privacy preferences; employees and/or applications try to access data for specific purposes; data subject changes their consent/privacy preferences; personal data is disclosed to a third party.</td>
<td>XML for structuring data; MongoDB for storing data.</td>
</tr>
<tr>
<td>ADvoCATE</td>
<td>2015-2019</td>
<td>Consent management in IoT environment.</td>
<td>Data protection ontology by Bartolini et al. [40]; Ethereum blockchain to maintain consent integrity and versioning.</td>
</tr>
<tr>
<td>SPECIAL-K</td>
<td>2017-2019</td>
<td>Consent for municipality road layout optimisation; sending bank travel insurance; sending traffic condition warning.</td>
<td>SPLog ontology modelling consent; MongoDB for storing data.</td>
</tr>
<tr>
<td>Davari et al.</td>
<td>2019</td>
<td>Management of consent and smart contracts with blockchain technology when Multi-National Companies (MNC) are involved.</td>
<td>XACML based access control model for implementing privacy framework; Blockchain framework Hyperledger Fabric for smart contract; PROV-O ontology for modelling consent according to GDPR; MongoDB for storing data.</td>
</tr>
<tr>
<td>CampaNeo</td>
<td>2019-2022</td>
<td>Consent for vehicle sensor data sharing.</td>
<td>Knowledge graphs for data modelling; CampaNeo ontology to define the concepts of campaign, data, processing, third-party entities, users and consent; GraphQL as an access point and schema for data.</td>
</tr>
<tr>
<td>Jaiman et al.</td>
<td>2020</td>
<td>Individual consent model for health data sharing platforms.</td>
<td>Data Use Ontology (DUO) for modelling consent and describing data use conditions; Discovery and Access Metric (ADA-M) ontology for classifying data use conditions and permissions; Ethereum blockchain for smart contract using the Solidity language.</td>
</tr>
<tr>
<td>Mahindrakar et al.</td>
<td>2020</td>
<td>GDPR compliance in real time; enforce data privacy policy when data is shared with third parties.</td>
<td>Privacy policy ontology for consent representation; GDPR ontology for GDPR articles; Ethereum private blockchain network - Ganache-CLI for smart contract; natural language processing for extracting privacy policies; AES encryption for encrypting data files.</td>
</tr>
<tr>
<td>smashHit</td>
<td>2020-2022</td>
<td>Consent for sensor data sharing in a smart city and for insurance purposes.</td>
<td>Ontologies for contract modelling; knowledge graphs for storing data about users, consent and contracts.</td>
</tr>
</tbody>
</table>

3.3.9. Summary

We summarise the overviewed research (completed and ongoing) from this section in Table 6. Looking back at the scope and main goal for each research project, it becomes clear that consent management is a complex multi-action process that is closely connected to the fields of data privacy and security.

Table 6 shows the overviewed solutions for consent management. Most of the projects and studies make use of semantic technology, namely ontologies and knowledge graphs, showing semantic technology as helpful data models for consent due to their ability to represent relationships between concepts. The projects SPECIAL-K [21], CampaNeo and studies by Rantos et al. [39], Jaiman et al. [42], Davari et al. [20], Mahindrakar et al. [44] using ontologies and knowledge graphs have demonstrated the value of semantic technology, namely knowledge graphs and ontologies for consent management. Further, considering the advantage of semantic technology, new projects like smashHit are also making use of ontologies and knowledge graphs for consent management. In addition to knowledge graphs and ontologies, studies like [20, 39, 42, 44] also make use of blockchain technology. The use of blockchain technology is adding value due to its ability to provide traceability and automatic code execution using a smart contract. In particular, the smart contract was used for executing the task of consent verification.

However, the research by Davari et al. [20] and Mahindrakar et al. [44] highlights the limitation that arises with the use of blockchain for storing data. The limitation is because of the immutability nature of...
the blockchain, which contradicts the user rights such as "the right to be forgotten" whenever the data subject revokes the consent. To deal with limitations due to immutability of the blockchain, external storage like a relational database, the file system is used for storing the data, and only the hashes are stored in the blockchain.

3.4. Standardisation Initiatives and Efforts

This section presents the current status of standards and standardisation efforts related to consent, namely Consent Receipts v1.1 [46], ISO/IEC 29184:2020 [50] and IAB Transparency and Control Framework [51].

3.4.1. Consent Receipts v1.1

The Consent Receipt v1.1 specification [46], published in 2018, provides an interoperable and transparent "record" of consent similar to a receipt after payment/sale of goods - for benefit to both Data Controllers and individual. The specification uses terms and definitions from ISO 29100:2011 [53] to describe consent, purposes, organisations, and recipients, and is structured as a flat-list or non-hierarchical schema with an implementation using JSON which adopters must implement for conformance. It lacks the necessary fields to represent and conform with requirements from recent laws such as GDPR. However, it provides a useful direction for creating and maintaining shared documentation for representation of consent that can be utilised by both the individual and controllers.

There is work underway to update the Consent Receipt with the recent developments and requirements, such as for GDPR. For this, Kantara has initiated the Advanced Notice & Consent Receipts Working Group [54] (ANCR). ISO/IEC have similarly initiated work on a new standard - ISO/IEC 27560 [55] Consent Record Information Structure.

3.4.2. ISO/IEC 29184:2020

ISO/IEC 29184:2020 [50] standard, published recently in June 2020, concerns the provision of privacy notices and requesting consent in an online context. It specifies requirements for information provided in a notice, its form and manner for comprehension, and role in validity of consent. It also dictates the process for the collection of consent in order for it to be valid. The standard notably raises the requirement of consent to be ‘explicit’ as the default, specifies risk assessment information, and advocates privacy and individual centric measures in both notice and consent related information and processes. 29184 specifically acknowledges the role of semantics and machine-readability for consent requests and records, and uses the Consent Receipt [46] specification as an example.

3.4.3. IAB Transparency and Control Framework

The Interactive Advertising Bureau (IAB) [51] is a non-profit organisation that creates and maintains standards for use within the online advertising network that involves some of the largest data operators and consent framework providers such as Google, Oracle, Adobe, Quantcast, OneTrust. Its "Trans-
Transparency and Control Framework’ (TCF) specification provides a protocol and data model for representing collected consent and its use within the online marketplace for ads based on the Real-Time Bidding (RTB) process. TCF consists of a controlled list of purposes, recipients, third-parties for data sharing, and controls associated personal data and based on use of legitimate interest and consent.

3.4.4. Summary
The standards and standardisation regarding consent is notably limited in terms of practical usage to IAB’s TCF framework. It is currently unclear what role such standards play in legal compliance, and their validity in different use-cases. However, the publication of ISO/IEC 29184, its acknowledgement of semantics and machine-readability for interoperable consent records, and the renewed interest in interoperable and machine-readable Consent Receipts shows promising developments in the future. This provides further motivation for inclusion of semantics in the consent management process based on these standards and their modeling of proposes and use-cases.

4. Best Practices and Recommendations
On the basis of the surveyed literature, this section is divided into subsections that present best practices for each of the four stages of the consent life-cycle (Figure 1) - request, comprehension, decision and use. The best practices are to provide guidelines on the ways to implement consent in organisations, as well as an input to researchers and policy makers on the possible future research. The following recommendations focus on the semantic and technical aspects of consent implementation, while considering standards (see Section 3.4), ethics and law (i.e. GDPR).

Before making specific recommendations, we would like to highlight that GDPR is just one of the many laws aimed at user’s privacy and rights. In Europe, for example, before the GDPR, the ePrivacy Directive was (and still is) one of the laws for personal data processing and privacy protection. ePrivacy and its derivative laws require consent for cookies, which is often combined with consent for personal data processing. In addition, each country has its own laws related to the matter. Reviewing them is not in the scope of this paper, however, we list several laws that one might want to consider. For example, Austria’s Data Protection Law (DSG) and Germany’s Federal Data Protection Act (BDSG) in Europe. Examples of laws regarding data privacy outside the EU are California’s Consumer Privacy Act (CCPA), The Notifiable Data Breach (NDB) in Australia, Brazil’s Lei Geral de Proteção de Dados (LGPD).

4.1. Request of Consent
Requesting consent can be seen as one of the most important stages in the consent life-cycle (Figure 1) as it defines whether or not data processing can begin. A successful consent request, which we view as

56 https://iabeurope.eu/transparency-consent-framework/
57 https://www.iab.com/guidelines/openrtb/
58 https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058
60 https://www.gesetze-im-internet.de/englisch_bdsig/
61 https://oag.ca.gov/privacy/ccpa
63 https://gdpr.eu/gdpr-vs-lgpd/
one that results in receiving individual’s consent, should be GDPR compliant. Having a semantic model for consent, which represents GDPR information in both human-readable and machine-readable format, would be beneficial to any system. Such model can be build with ontologies as shown in Section 3.1. However, consent requests are made to the user thus a visualisation of the request itself is needed as well. Further, once requested and given by the individual the consent needs to be managed, for example, to check who is allowed or denied to do some activity on what data.

4.2. Comprehension of Consent

Semantic technology helps achieve a common understanding between multiple entities by representing information in both human-readable and machine-readable formats. For a machine, representing the
### Recommendations for the Comprehension of Consent

<table>
<thead>
<tr>
<th>Recommendations</th>
<th>Relevant Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semantic Models for Consent</strong></td>
<td><strong>Section 3.1</strong></td>
</tr>
<tr>
<td>• Understand the domain. An ontology would reflect the ontology engineer’s understanding of a specific domain. Begin by selecting an ontology engineering methodology e.g. of Noy and McGuinness [48]. We recommend deriving all important concepts and how they might be related. Once this is clear one can translate the knowledge into an ontology by following different methodologies as presented in.</td>
<td></td>
</tr>
<tr>
<td>• Select an ontology language based on the desired functionality. Most of the consent ontologies in Section 3.1 are built with OWL. In comparison to OWL, OWL2 offers more expressivity by allowing the use of keys, property chains qualified cardinality restrictions, richer data ranges, asymmetric, reflexive, disjoint properties, and enhanced annotation capabilities. Other languages such as RDFs, KIF and DAML+OIL, and popular upper level ontologies such as Dublin Core can be used as well. For example, a combination of several ontology syntaxes is possible as well. The Colpri [22] ontology is built with both OWL and SKOS. A detailed comparison of ontology languages is presented in [49].</td>
<td></td>
</tr>
</tbody>
</table>

| **Consent Visualisation**                                                         | **Section 3.2**   |
| • Use graphical visualisations to represent the data flow. For example, graphs can be easier to understand by humans than text, as they provide a visualisation of the main entities and the connections between them. The graphical visualisations in the overviewed tools have proven to be useful and to provide individuals with the information in an easily comprehensible way. |                     |
| • Include the end-user. In the Data Track tool [29], the end user is visualised at the centre of the graph. This has resulted in individuals feeling more involved and interested in what is happening to their data. |                     |
| • Allow interactivity. The Data Track tool [32], CoRe [34] and CURE [35] UIs and the Privacy Dashboard [33] have all included interactive elements in their visualisations. For example, Data Track allows individuals to explore the provided graphical visualisation by expanding and collapsing certain UI fields and the graph itself. CoRe and CURE both allow interactivity when individuals give consent - one can select for what purpose to give consent and to follow the data flow for that purpose. |                     |
| • Accessibility. Individuals should be able to understand what is presented and also be able to interact with it directly. Individuals with disabilities should be considered as well. For example, developing interfaces that recognise one’s speech and also allow dictation of text and similar features (e.g. in the MAC iOS operating system) would be beneficial for individuals who suffer from blindness. |                     |

| **Consent Management**                                                           | **Section 3.1, Section 3.3** |
| • Use semantic technology. Consent management can be performed automatically by any machine at any time, however, without semantics a machine simply executes commands specified by an individual and yields a result. It does not actually understand what the data or the commands mean. Semantic technology changes this as it adds value to things and helps machines become aware. By enhancing machines with semantics one would be able to climb higher in the so-called DIKW (data, information, knowledge, wisdom) [50] hierarchy and reach the knowledge level. |                     |

### 4.3. Decision about Consent

When it comes to giving consent, the decision rests in the hands of the user. All people are biased in their own way due to their upbringing and current environment. While some users might give consent just to be “done” with the process, the choice of others could be affected by many factors such as the information that is presented, the level of detail, specific interface design [29]. By reviewing existing concepts with languages such as OWL or RDF is enough, however, this is not the case with end users.

End-users have different needs and understanding of information. Further, one’s knowledge of the semantic web could also be a challenge thus a simple yet effective visualisation of consent is needed. This visualisation is directly linked to GDPR’s consent requirement regarding requesting consent (Section 1). Humans are visual creatures thus a visualisation of the required data would be more efficient in comparison to presenting one with long privacy policies written in legal jargon. In this section we provide guidelines (Table 8) for visualising information to end-users based on the reviewed literature (Section 3.2) in the area of consent visualisation for improving comprehension. In addition, we present recommendations (Table 8) on how to enhance a machine’s understanding of things with semantic technology. The recommendations are divided into three sections: semantic model for consent, consent visualisation and consent management, all of which relate to the comprehension of consent.
4.4. Use of Consent

User’s consent can be used in many ways (e.g. compliance checking, reasoning, as a proof of contract) and each way requires different system functionalities. All these actions performed with consent, could be summarised as consent management (see Section 3.3). The recommendations in Table 10 are divided into three sections: semantic model for consent, consent visualisation and consent management, all of which relate to the use of consent.
5. Conclusions

Semantic technology such as ontologies are the key to achieving a common understanding between machines and humans. Although they have been around for many years, there is much more to discover about their possible applications in different fields. For example, understanding the benefit of semantics in the law domain, which we address by specifically looking at semantic technology for consent implementation according to GDPR.
In this paper we presented an overview of existing semantic solutions for implementing consent and recommendations for implementing consent with semantic technology. To be specific, we provided guidelines for building a semantic model for consent, graphically visualising consent to individuals for better comprehension and for consent management.

As we have shown with the overviewed work, it is possible and useful to have a semantic model for consent in the form of an ontology that models consent through its whole life-cycle (Figure 1). For the request of consent, a semantic model provides a description of all the information required by laws (e.g. GDPR) for informed consent, thus it provides a common understanding of the law requirements that both machines and humans understand and need to follow. Based on the underlying semantics a machine is able to create the links between the consent decision and all information related to it. During the comprehension step, the semantic model helps to translate the human knowledge into machine knowledge and to establish a common understanding of the meaning of consent, the risks and consequences associated with it to other humans. An ontology can also model different states of consent, for example consent revocation and the rules that apply in such situation so that a machine is able to handle the consent state change in compliance with the law and most of all in a meaningful way. Finally, the use of consent or also called in this paper "consent management", benefits from the traceability, transparency and faster and easier knowledge discovery that a semantic model offers.

All of these semantic model capabilities can be utilised when actions such as consent validation and compliance checking need to be performed. Although a semantic model offers many advantages, the difficulty of implementing informed consent is still present due to the need for one to not only understand and model laws such as GDPR, but to also integrate them with suitable technologies (e.g. blockchain is not a suitable storage for informed consent as defined by GDPR [20][44]). Further, complex issues regarding consent that need to be addressed are traceability and compliance checking.

As mentioned in Section 4, the jurisdictional limitation of laws that means there are several relevant laws that regulate consent in relation with data processing - applicable within their own jurisdiction or domain. For example, the California Consumer Privacy Act (CCPA)61 (effective since January 2020) applies to companies in the state of California, USA and is consumer and privacy oriented as compared to GDPR’s focus on data protection. Ontologies and semantics in general can help organisations to identify and address common requirements across such laws, for example similarities between GDPR and CCPA [61]. The challenge for such approaches lies with the law-specific terms and requirements, such as the notion of ‘do-not-sell’ under CCPA which permits individuals to opt-out of data sharing (termed ‘selling’ under CCPA) to third parties. One possible solution for this could be to utilise a common ontological framework and build extensions for specific legal requirements - such as the approach taken by the Data Privacy Vocabulary (DPV) vocabulary.

In conclusion, this survey paper focused mainly on ontologies as a semantic model for consent and how they could be used for consent management. The evolution of the models and techniques built on them will include semantic models such as schemas that have been used for many years already, as well as newer solutions built with knowledge graphs [62], addressing the desired systems’ functionalities.

6. Acknowledgements

This research has been supported by the smashHit European Union project funded under Horizon 2020 Grant 871477. Harshvardhan J. Pandit is funded by the Irish Research Council Government of Ireland Postdoctoral Fellowship Grant GOIPD/2020/790, by European Union’s Horizon 2020 research...
and innovation programme under NGI TRUST Grant 825618 for Privacy as Expected: Consent Gateway project, and through the ADAPT SFI Centre for Digital Media Technology which is funded by Science Foundation Ireland through the SFI Research Centres Programme and is co-funded under the European Regional Development Fund (ERDF) through Grant 13/RC/2106_P2.

References


