Access Control and the Resource Description Framework: A Survey

Sabrina Kirrane\textsuperscript{a}, Alessandra Mileo\textsuperscript{b}, Stefan Decker\textsuperscript{c}

\textsuperscript{a} Vienna University of Business and Economics, Austria
E-mail: sabrina.kirrane@wu.ac.at
\textsuperscript{b} Insight Centre for Data Analytics, National University of Ireland, Galway, Ireland
E-mail: alessandra.mileo@insight-centre.org
\textsuperscript{c} RWTH Aachen University, Germany
E-mail: stefan@stefandecker.org

Abstract. In recent years we have seen significant advances in the technology used to both publish and consume structured data using the existing web infrastructure, commonly referred to as the Linked Data Web. However, in order to support the next generation of e-business applications on top of Linked Data suitable forms of access control need to be put in place. This paper provides an overview of the various access control models, standards and policy languages, and the different access control enforcement strategies for the Resource Description Framework (the data model underpinning the Linked Data Web). A set of access control requirements that can be used to categorise existing access control strategies is proposed and a number of challenges that still need to be overcome are identified.

Keywords: Access Control, Policies, Linked Data, Semantic Web, Resource Description Framework

1. Introduction

The term Linked Data Web (LDW) is used to describe a World Wide Web where data is directly linked with other relevant data using machine-readable formats \cite{DBLP:journals/corr/abs-1602-04711, DBLP:conf/edาน/DeckerF10}. Although the technology underpinning the LDW has been in existence for a number of years, up until now data publishers have primarily focused on exposing and linking public data. With the advent of update languages for the Resource Description Framework (RDF) data model, such as SPARQL 1.1 \cite{DBLP:journals/corr/abs-1501-01613}, the LDW has the potential to evolve from a medium for publishing and linking public data, to a dynamic read/write distributed data source. Such an infrastructure will be capable of supporting not only data integration of public and private data, but also the next generation of electronic business applications. However, in order to make the move from simply linking public data to using the Linked Data infrastructure as a global data space, suitable security mechanisms need to be put in place.

Security in general and access control in particular have been extensively studied by both the database and the information system communities, among others. Early work on access control policy specification and enforcement within the Semantic Web community focused on: representing existing access control models and standards using semantic technology; proposing new access control models suitable for open, heterogeneous and distributed environments; and devising languages and frameworks that can be used to facilitate access control specification and maintenance. Later researchers examined access control for the RDF data model in general and access control propagation,
2. Access Control Models and Standards

Generally speaking the term access control is used to refer to the model, which is the blueprint that is used to guide the access control process; the policy language, which defines both the syntax and the semantics of the access control rules; and the framework, which is a combination of the access control model, the language and the enforcement mechanism. At its most basic, an access control rule (otherwise known as an authorisation) can be represented as a tuple \( \langle S, R, AR \rangle \) where \( S \) denotes the subject (entities requesting access to resources), \( R \) denotes the resource (entities to be protected) and \( AR \) represents the access rights (permissions and prohibitions often based on actions pertaining to the resource). Sets of access control rules are collectively referred to as an access control policy. The decision to grant or deny access is based on two distinct processes, authentication and authorisation. Authentication involves the verification of credentials (you are who you say you are). Whereas, authorisation is the process of granting or denying access to system resources based on credentials.

This section describes both well known and emerging access control models, and relevant standardisation efforts. In each instance, a description of the access control model/standard is provided, along with details of how the model/standard has been applied to or enhanced using Semantic Web technologies.

2.1. Access Control Models

Mandatory Access Control (MAC), Discretionary Access Control (DAC) and Role Based Access Control (RBAC) are the most prominent access control models found in the literature, and used in practice. View Based Access Control (VBAC) is a complementary access control model which grants access to sets of entities, logically structured as views. More recently researchers have proposed new access control models, that are deemed more suitable for the open, heterogeneous and distributed architecture of the web. Primary research efforts, involve using properties (relating to the subject, resource or the environment) as opposed to identities, in order to determine if access to resources should be permitted. Attribute Based Access Control (ABAC) and Context Based Access Control (CBAC) are the predominant works in this area.

In this section we present the vocabularies used to specify the access control model and provide details of the different enforcement mechanisms. A summary of the existing proposals is presented in Table 1 and the core ideas are represented on a timeline in Figure 1.

2.1.1. Mandatory Access Control

MAC limits access to resources using access control policies determined by a central authority [79]. The central authority is responsible for classifying both subjects and resources according to security levels. Resources are assigned labels that represent the security level required to access the resource, and only subjects with the same security level or higher are granted access. MAC was originally developed for military ap-
Table 1

<table>
<thead>
<tr>
<th>Model</th>
<th>Policy Representation</th>
<th>Enforcement Mechanism</th>
<th>Enforcement Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAC</strong></td>
<td>Kodali et al. [54]</td>
<td>MAC concepts and security levels modelled in DAML+OIL</td>
<td>DL &amp; DQL query based enforcement</td>
</tr>
<tr>
<td>Yagüe del Valle et al. [104]</td>
<td>subject, resource and policy semantic metamodels</td>
<td>XML based enforcement</td>
<td>attribute based access control</td>
</tr>
<tr>
<td><strong>DAC</strong></td>
<td>Gabillon and Letouzey [33]</td>
<td>user, resources or the environment</td>
<td>rule based enforcement</td>
</tr>
<tr>
<td>Kirrane et al. [52]</td>
<td>GRANT &amp; REVOKE SPARQL queries</td>
<td>rule based enforcement</td>
<td>DAC requirements</td>
</tr>
<tr>
<td>Kodali et al. [54]</td>
<td>DAC concepts and subjects modelled in DAML+OIL</td>
<td>DL &amp; DQL query based enforcement</td>
<td>multiple AC models</td>
</tr>
<tr>
<td>Yagüe del Valle et al. [104]</td>
<td>subject, resource and policy semantic metamodels</td>
<td>XML based enforcement</td>
<td>attribute based access control</td>
</tr>
<tr>
<td><strong>RBAC</strong></td>
<td>Alcaraz Calero et al. [2]</td>
<td>roles mapped to CIM and modelled in OWL</td>
<td>DL &amp; rule based enforcement</td>
</tr>
<tr>
<td>Wu et al. [103]</td>
<td>roles as classes in OWL</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ferrini and Bertino [26]</td>
<td>roles modelled as XACML policies &amp; RBAC constraints modelled in OWL</td>
<td>XACML and DL based enforcement</td>
<td>roles &amp; constraints in OWL</td>
</tr>
<tr>
<td>Finnin et al. [38]</td>
<td>roles as classes in OWL &amp; roles as instances in OWL</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kodali et al. [54]</td>
<td>RBAC concepts and roles modelled in DAML+OIL</td>
<td>DL &amp; DQL query based enforcement</td>
<td>multiple AC models</td>
</tr>
<tr>
<td>Yagüe del Valle et al. [104]</td>
<td>subject, resource and policy semantic metamodels</td>
<td>XML based enforcement</td>
<td>attribute based access control</td>
</tr>
<tr>
<td><strong>VBAC</strong></td>
<td>Dietzold and Auer [25]</td>
<td>triples, classes and properties</td>
<td>rule &amp; SPARQL based enforcement</td>
</tr>
<tr>
<td>Gabillon and Letouzey [33]</td>
<td>user, resources or the environment</td>
<td>rule &amp; SPARQL based enforcement</td>
<td>SPARQL CONSTRUCT and DESCRIBE queries &amp; named graphs</td>
</tr>
<tr>
<td>Li and Cheung [57]</td>
<td>concepts, relations and filter conditions</td>
<td>rule based enforcement</td>
<td>data filtering</td>
</tr>
<tr>
<td>Mühliesen et al. [29]</td>
<td>triple patterns, resources or instances policies</td>
<td>rule based enforcement</td>
<td>using SWRL rules to generate temporary named graph</td>
</tr>
<tr>
<td><strong>ABAC</strong></td>
<td>Cirio et al. [15]</td>
<td>ABAC modelled in OWL-DL</td>
<td>DL and SPARQL based enforcement</td>
</tr>
<tr>
<td>Priebe et al. [69]</td>
<td>-</td>
<td>-</td>
<td>metadata based access control pattern</td>
</tr>
<tr>
<td>Priebe et al. [67]</td>
<td>XACML modelled in OWL</td>
<td>XACML based enforcement</td>
<td>XACML extended with OWL reasoning</td>
</tr>
<tr>
<td>Sternekil et al. [85]</td>
<td>RDF properties used to specify subjects and objects</td>
<td>-</td>
<td>policy decision and enforcement points</td>
</tr>
<tr>
<td><strong>CBAC</strong></td>
<td>Corradi et al. [16]</td>
<td>context ontology</td>
<td>high level policy enforcement model</td>
</tr>
<tr>
<td>Montanari et al. [59]</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Costabile et al. [17]</td>
<td>context specified using name/value pairs</td>
<td>Java &amp; SPARQL based enforcement</td>
<td>using ASK queries to check policies</td>
</tr>
<tr>
<td>Shei and Cheng [83]</td>
<td>context and policy ontologies</td>
<td>SWRL rule based enforcement</td>
<td>-</td>
</tr>
</tbody>
</table>

Applications and therefore it is best suited to closed environments, where a great deal of control is required [6]. Given the open, heterogeneous and distributed nature of the web, it is not surprising that MAC has not gained much traction among Semantic Web researchers. Primary research efforts focus on: (i) defining vocabularies that can be used to support multiple access control models [54]; and (ii) using attributes to represent different credentials (e.g. identities, roles and labels) [104]. A summary of the various policy specification and enforcement mechanisms are presented below:

**Policy Specification.** Kodali et al. [54] propose a DAML+OIL ontology and demonstrate via example how it can be used to represent MAC, DAC and RBAC policies. In the case of MAC both the subject and resources are assigned security levels. A subject is granted access a to resource if the their security level dominates the security level of the resource. One of the drawbacks of the approach is that a resource can only be governed by one access control model (either MAC, DAC or RBAC).

An alternative strategy is proposed by Yagüe del Valle et al. [104]. Like Kodali et al. [54], the authors provide support for multiple access control models, however rather than defining an ontology the authors propose an ABAC model and discuss how it can be used to support not only MAC but also DAC and

1DAML+OIL, http://www.w3.org/TR/daml+oil-reference
RBAC.

**Enforcement Framework.** Kodali et al. [54] describe a unifying framework which can be used to enforce a number of different access control models, MAC being one of them. The authors use DQL\(^2\) queries (i.e. DAML+OIL query patterns) together with a DAML-JessKB Description Logic (DL) reasoner in order to create an access restricted view of a multimedia document for a particular security level.

Although Yagüe del Valle et al. [104] indicate that they adopt an XML based enforcement mechanism for their attribute based access control policies, they do not describe the enforcement framework.

### 2.1.2. Discretionary Access Control

DAC policies associate one or more subjects with sets of access rights pertaining to one or more resources. Like MAC, DAC restricts access by means of a central access control policy, however users are allowed to override the central policy and can pass their access rights on to other users, a process known as delegation [81]. According to Weitzner et al. [102], the web needs discretionary, rule based access control. Although the authors describe a motivating scenario and present a potential solution, they focus primarily on the general architecture of the system, as opposed to investigating how discretionary access control can be modelled or enforced. When it comes to intersection between DAC and RDF, research efforts to date have focused on: (i) supporting multiple access control models [104,54]; (ii) providing support for the delegated of access rights to others [33,35,52]. A summary of the various policy specification and enforcement mechanisms are presented below:

**Policy Specification.** The ontology proposed by Kodali et al. [54] and the ABAC model proposed by Yagüe del Valle et al. [104] can also be used to specify DAC policies. In both instances, the authors focus on the specification of policies rather than the delegation of permissions. In contrast, Kirrane et al. [53] provide a summary of discretionary access control requirements for the RDF data model, based on the different characteristics of the graph data model compared to relational and tree data models. The authors

---

\(^2\)DQL, http://www.daml.org/dql/
suggest extending the SPARQL query language with two new commands `GRANT` and `REVOKE` that could be used to manage the delegation of access rights in RDF databases. In follow-up work Kirrane et al. [52] discuss how DAC can be used together with declarative policies. In particular, the authors identify the need for an additional authorisation element that is used to represent the person that generated the authorisation.

**Enforcement Framework.** According to Kodali et al. [54] and Yagüe del Valle et al. [104] their general access control frameworks can also be used to enforce DAC policies, however as mentioned previously the authors do not consider the delegation and revocation of permissions.

In contrast both Gabillon and Letouzey [33] and Kirrane et al. [52] focus specifically on access control delegation. Gabillon and Letouzey [33] allow users to define security policies for RDF graphs and SPARQL views (SPARQL `CONSTRUCT` and `DESCRIBE` queries) that they own. Users may delegate rights to other users by specifying an authorisation which grants `CONSTRUCT` and `DESCRIBE` privileges to their RDF dataset, one or more RDF graphs or views of the RDF dataset. Although the authors describe the delegation process no consideration is given to the revocation of policies. Kirrane et al. [52] also propose an ownership model, whereby the data producer is granted full access to the data items they create. Although they present an administration algorithm that can be used to generate new authorisations, they do not provide algorithms for the delegation and revocation processes. The authors state that neither the grant nor the revoke algorithms are dependent on the data model and as such traditional revocation approaches such as cascading and non-cascading could be used in conjunction with the proposed framework.

### 2.1.3. Role Based Access Control

RBAC restricts access to resources to groups of users, with common responsibilities or tasks, generally referred to as roles. In RBAC, users are assigned to appropriate roles and access to resources is granted to one or more roles, as opposed to users directly [82]. The term session is commonly used to refer to the set of roles that the user is currently assuming (their active roles). Whereas, role deactivation is generally used to refer to the process whereby a user is removed from a role. Depending on the usecase, roles may be organised to form either a hierarchy or a partial order. Such structures are used to simplify access control specification and maintenance. Constraints are commonly used to enforce conditions over access control policies. For RBAC, these constraints take the form of both static and dynamic separation of duty (a user cannot be assigned to two roles simultaneously) and prerequisites (a user can only be assigned to a role if they have already been assigned another required role).

When it comes to RBAC research efforts have primarily focused on: (i) modelling RBAC entities as classes [54, 103, 30] or instances [30]; and (ii) adapting existing vocabularies, such as the eXtensible Access Control Markup Language (XACML) [28] and the Common Information Model (CIM) [2], in order to cater for RBAC. An overview of the various RBAC policy specification and enforcement mechanisms are presented below:

**Policy Specification.** Kodali et al. [54] propose an ontology that can be used to associate access control policies with roles. Wu et al. [103] provide a basic modelling for RBAC concepts and constraints using OWL. User, Role, Permission and Session entities are represented as classes. While, the following properties are used to represent relationships:

- `hasRole` (assigns a user to a role);
- `hasPermission` (associates a role with a permission);
- `belongsTo` (maps a session to a single user); and
- `hasActiveRole` (maps a set of roles to a session).

Two additional properties are used to model separation of duty and prerequisite constraints:

- `conflictRole` (indicates that there is a conflict between two roles); and
- `prerequisitesRole` (specifies that one role is dependent on another).

Finin et al. [30] build on the work proposed by Wu et al. [103] by examining the advantages and disadvantages of representing roles as classes and roles as instances. When roles are represented as instances the modelling is simple and more concise. Whereas, when roles are represented as classes, it is possible to determine subsumption relationships according to a users active role and the role hierarchy, using standard description logic subsumption. In order to cater for role deactivation it should be possible to temporarily remove a user from a role. However, as OWL is mono-

---

3OWL, [www.w3.org/TR/owl-ref/](http://www.w3.org/TR/owl-ref/)
tonic "state changes such as role deactivations, and modifying role permission assignments must be handled outside the reasoners" \[30\]. For example, assuming that a user has a set of roles assigned to them. When the user interacts with the system, it should be possible for them to choose to activate (or deactivate) any of their roles. The activated roles determine which permissions are available to the user at a given time. As such, the authors use N3Logic rules in order to cater for separation of duty and role deactivation constraints.

Rather than propose a new vocabulary, Ferrini and Bertino \[28\] and Alcaraz Calero et al. \[2\] demonstrate how existing vocabularies can be modelled in OWL. Ferrini and Bertino \[28\] discuss how the eXtensible Access Control Markup Language (XACML) \[4\], an attribute based access control language and framework, proposed by the Advanced Open Standards for the Information Society (OASIS), and OWL can be used to specify and enforce RBAC. While, Alcaraz Calero et al. \[2\] demonstrate how the Common Information Model (CIM) \[5\] (a standard vocabulary used to represent information technology objects and the relationship between them) can be used to represent RBAC. The authors provide a mapping between RBAC concepts and the CIM vocabulary which they model using OWL.

**Enforcement Framework.** According to Yagüe del Valle et al. \[104\] their general access control frameworks can also be used to enforce RBAC policies. However, as stated previously Yagüe del Valle et al. \[104\] do not describe the enforcement framework. In contrast, Ferrini and Bertino \[28\] discuss how ABAC can be used enforce RBAC policies. In the proposed framework the user submits their role as a subject attribute. On receipt of the request the system extracts the role from the subject attribute and the description logic reasoner is used to retrieve additional roles that can be inferred from the OWL ontology. These roles are subsequently fed into the XACML engine. Finally the XACML engine consults the XACML policy in order to determine if access should be granted. As the policies are not modelled in OWL it is possible to support role deactivation. However, with the existing modelling it is not possible to exploit reasoning over policy resources or access rights.

---

*\[XACML\](https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml)*  
*\[CIM\](http://www.dmtf.org/standards/cim)*  
In the architecture proposed by Alcaraz Calero et al. \[2\] constraints are represented as rules using the Semantic Web Rule Language (SWRL) \[6\]. For example, in order to model a separation of duty constraint, a rule is defined which checks if there are any common instances between roles and if so generates an exception. The authors describe how a SPARQL enabled reasoner, such as Pellet, can be used to query the OWL policy, in order to determine if access should be granted or denied. In Section \[5\] we will discuss the benefits and limitations of both ontology and rule based enforcement strategies.

### 2.1.4. View Based Access Control

VBAC \[57\] is used in relational databases to simultaneously grant access to one or more relations, tuples, attributes or values. A similar approach is used in Object Oriented Access Control (OBAC) \[27\], where access rights are granted to sets of application objects. Primary research efforts with respect to VBAC and RDF focus on: (i) using rules to grant and deny access to views of the data \[57,60\]; and (ii) using SPARQL `CONSTRUCT` and `DESCRIBE` queries to grant and deny access at multiple levels of granularity \[25,33\].

A summary of the various VBAC policy specification and enforcement mechanisms are presented below:

**Policy Specification.** Li and Cheung \[57\] allow authorisations to be specified using a combination of ontology concepts, relations and filter conditions (equal to, less than and greater than) and use rules to generate access restricted views of the dataset. Similarly, Mühleisen et al. \[60\] allow access control policies to be specified for *triple patterns, resources or instances* using SWRL rules. In the case of Li and Cheung \[57\] access control policies are associated with a particular *subject*, whereas in the case of Mühleisen et al. \[60\] access control policies are specified using contextual information pertaining to the *user, resources or the environment*.

Both Dietzold and Auer \[25\] and Gabillon and Letouzey \[33\] use a combination of rules and filters defined using SPARQL queries in order to construct a access restricted view of the data. Dietzold and Auer \[25\] propose access control policy specification at multiple levels of granularity (triples, classes and properties). Authorisations are used to associate filters (SPARQL `CONSTRUCT` queries) with users and resources. Gabillon and Letouzey \[33\] extend the modelling proposed by Dietzold and Auer \[25\] by proposing an access con-

---

*\[SWRL\](http://www.w3.org/Submission/SWRL/)*
trol model which can be used to grant/deny access to both SPARQL construct and describe queries and named graphs.

**Enforcement Framework.** Both Li and Cheung [57] and Mühleisen et al. [60] discuss how rules can be used to generate a subset of the data that the requester is authorised to access. In the case of Li and Cheung [57] when a requester submits a query, the system uses the rules to generate a view of the authorised data. While Mühleisen et al. [60] use rules to generate a temporary named graph containing only authorised data. The requesters query is subsequently executed against the temporary named graph and the results are returned to the user.

Both Dietzold and Auer [25] and Gabillon and Letouzey [33] describe how RDF data can be logically organised into views using SPARQL construct and describe queries. When a requester submits a query, a virtual model is generated based on the matched authorisations. The query is subsequently executed against the virtual model, which only contains data that the requester is authorised to access.

### 2.1.5. Attribute Based Access Control

ABAC was designed for distributed systems, where the subject may not be known to the system, prior to the submission of a request. ABAC grants or denies access to resources, based on properties of the subject and/or the resource, known as attributes. Primary research efforts to date focused on: (i) using rules to grant access based on attributes directly or indirectly using roles [85]; (ii) using ontologies to grant access based on attributes directly or indirectly using roles [15]; and (iii) proposing a pattern which can be used to guide the development of attribute based access control frameworks [66]. Primary research efforts with respect to ABAC policy specification and enforcement mechanisms are presented below:

**Policy Specification.** Priebe et al. [67], discuss how attribute based access control policies specified using XACML can be modelled using OWL. While, Cirio et al. [15] demonstrate how OWL-DL can be used to express role based access control concepts and extend the modelling to cater for attributes. The authors use rdfs:domain and rdfs:range properties to assert knowledge instead of using them as constraints. Such an approach simplifies policy specification as concepts can be defined implicitly. In the presented modelling policies are specified for classes as opposed to instances. For example, it is possible to state that a student prepares a thesis and is advised by an advisor; and an advisor advises a student and reviews a thesis. However, it is not possible to say that an actual thesis prepared by a particular student is the same thesis instance that is retrieved by the respective advisor. As such, the authors use two predicates requiresTrue and requiresFalse (with domain Role and range String) in order to specify run time constraints based on user attributes. The constraints are specified using SPARQL queries, which are executed against the knowledge base at runtime. Similarly, Stermsek et al. [85] discuss how attributes can be used to specify access control directly using rules and indirectly using roles. In the former, the requesters attributes are compared against a policy, which indicates the attributes necessary to access a resource. Whereas in the latter, permissions are assigned to roles and the requesters attributes are used to determine the access rights or roles the subject should be mapped to.

**Enforcement Framework.** Priebe et al. [67], discuss how XACML policies together with a reasoning engine, allow for deductive reasoning based XACML policies modelled using OWL. Stermsek et al. [85] discuss three different mechanisms that can be used to obtain the attributes pertaining to a subject: (i) subjects can send all of their attributes to the server, with the initial request; (ii) the server can request particular attributes from the subject, once the initial request is submitted; and (iii) given the subject may be cautious about giving out the requested credentials to an unknown entity, both the server and the client could exchange policies and attributes (a process commonly known as trust negotiation). While, Cirio et al. [15] discuss how SPARQL ask queries can be used to represent constraints that are evaluated at runtime using a custom policy decision point, which wraps a description logic reasoner.

Priebe et al. [66], inspired by software design patterns, present a Metadata-Based Access Control (MBAC) pattern, which aims to encapsulate best practice with respect to attribute based access control. In the presented pattern subjects and objects are modelled as sets of attribute/value pairs. While, authorisation subjects and authorisation resources are described in terms of required attributes and corresponding values.
2.1.6. Context Based Access Control

CBAC uses properties, pertaining to users, resources and the environment, to grant/deny access to resources. In light of new and emerging human-computer interaction paradigms, such as ubiquitous computing and the internet of things, access control based on context has been gaining traction in recent years. Existing proposals for CBAC can be summarised as follows: (i) proposing a framework which differentiates between physical and logical context [16,59]; and (ii) using ontologies and rules to specify access control policies that take context relating to subjects, objects, transactions and the environment into consideration [83]; and (iii) using ontologies and SPARQL ASK queries to restrict access to named graphs [17]. A summary of the CBAC policy specification and enforcement mechanisms are presented below:

Policy Specification. Shen and Cheng [83] propose context and policy ontologies that are used to specify both positive and negative authorisations and obligations. Context provides a level of indirection between subjects and permissions. The authors propose four different types of context, that are relevant from a access control perspective: (i) Subject contexts (properties pertaining to a subject); (ii) Object contexts (information relating to resources); (iii) Transaction contexts (either current or past information relating to a particular action); and (iv) Environment contexts (other contextual information not relating directly to the subject, the resource or the action, for example time of day). An authorisation permits/prohibits an action, based on sets of contexts supplied by the user. Actions are used to represent operations that a subject wishes to perform and Permission Assignments are used to associate contexts with actions. Rules are used to insert new authorisations, based on contextual information, into a knowledge base. Like Shen and Cheng [83], Corradi et al. [16] and Montanari et al. [59] propose access control policies that are composed of mappings between context and permissions. Complex policies are generated using conjunction, disjunction and negation operators.

Costabello et al. [17] propose a policy language that can be used to restrict access to named graphs using contextual information supplied by the requester. An access control policy is a tuple \( (\text{ACS},\ AP,\ S,\ R,\ AEC) \), where \( \text{ACS} \) is a set of access conditions (specified using SPARQL ASK queries); \( AP \) is a set of access privileges (CREATE, READ, UPDATE or DELETE); \( S \) denotes the subjects to be protected; \( R \) represents the named graphs to be protected; and \( AEC \) is the evaluation context specified using name value pairs (verified using SPARQL BINDINGS).

Enforcement Framework. Corradi et al. [16] and Montanari et al. [59] propose a context based access control model and framework, called UbiCOSM. The proposed access control model uses context to group policies. The authors distinguish between physical and logical context. The former relates to the physical location denoted by geographical coordinates. Whereas, the latter refers to logical properties pertaining to the users and resources. When a user requests access to a resource, the UbiCOSM enforcement framework retrieves the relevant policies and generates a view based on the users permissions.

Shen and Cheng [83] propose a semantic-aware context-based access control (SCBAC) model and demonstrate how together ontologies and rules can be used to generate authorisations. In the proposed framework an access request is represented as a tuple \( (U, P, C, R) \) where user \( U \), requests privilege \( P \) on resource \( R \), in light of a given context \( C \). Access is enforced by representing access requests as SPARQL queries that are executed over the knowledge base. However, given OWL is monotonic, it is not clear how changes to contextual information are handled in the proposed approach.

In Costabello et al. [17], in addition to the SPARQL query that the user wishes to execute, the user provides their access credentials, in the form of a SPARQL UPDATE query, which contains contextual data. The enforcement framework stores the contextual data in a named graph and retrieves the authorisations that match the query type. In order to determine if access is permitted, the ASK query and the BINDINGS, that are specified in the authorisation, are executed against the users contextual graph. If the ASK query returns true then the query is rewritten to include the corresponding named graph.

2.2. Access Control Standardisations

In recent years, there have been a number of standardisation efforts, by both the World Wide Web Consortium (W3C) and the Organization for the Advancement of Structured Information Standards (OASIS), in relation to access control for web data. This section provides a high level overview of the relevant standards and details how they have been adapted or enhanced using semantic technology. A summary of the
existing work relating to each of the aforementioned standards is presented in Table 2 and the key concepts are represented in Figure 2 using a timeline. In each instance we make the distinction between simply applying versus extending the standard.

2.2.1. eXtensible Access Control Markup Language
The eXtensible Access Control Markup Language (XACML) is an OASIS standard, which is used to represent attribute based access control policies [71]. XML was chosen as the representation formalism for the policy language as it: (i) can be used to represent information in a human and machine readable manner; (ii) can easily be adapted to represent different access control requirements; and (iii) is widely supported by software vendors. The specification provides an XML schema which can be used to represent attribute based access control policies. The root of an XACML policy is a Policy or a PolicySet (used to represent multiple policies). Policies are composed of sets of Rules that are in turn composed of sets of Targets (conditions relating to Subjects, Resources and Actions) and an Access Decision (permit, deny or not applicable). A Request is represented as a tuple \((\text{subject}, \text{resource}, \text{action}, \text{environment})\). Where \text{subject} represents the entity requesting access, \text{resource} denotes the object to be protected, \text{action} defines the type of access and \text{environment} represents the requesters attributes. The XACML framework is composed of the following components:

- a policy decision point, which evaluates policies and returns a response;
- a policy enforcement point, which is responsible for making decision requests and enforcing the decisions;
- a policy information point, which obtains attributes pertaining to subjects, resources and the environment; and
- a policy administration point, which enables policies and sets of policies to be created and updated;

A number of researchers have used Semantic Web technologies to supplement the existing XACML framework. Primary research efforts include: (i) extending XACML to cater for deductive reasoning over attribute ontologies [67,68,14]; (ii) extending XACML policies with context [32]; and (iii) adapting XACML to work with roles [28].

Extensions. Priebe et al. [67,68] present an extension to XACML, which enables deductive reasoning over attribute ontologies, specified using OWL. In order to support reasoning, the authors propose two additional architecture components: an ontology administration point and an inference engine. Attribute ontologies are created and updated using the ontology administration point. If the attributes presented by the requester are not explicitly stated in the access control policy, the system attempts to infer the required access rights from the policy attributes, requester attributes and the attributes ontology, using the inference engine. Chen and Stuckenschmidt [14] also extend XACML with OWL deductive reasoning capabilities. XACML policies are specified using OWL and access control is enforced via query rewriting. The authors use SPARQL filters to both permit and deny access to instance data. Reasoning over the data, represented in the filters, is delegated to reasoners which support OWL and SPARQL.

Franzoni et al. [32] demonstrate how XACML can be extended to consider access control based on contextual properties pertaining to either the user or the application. In addition to standard access control policies, specified using XACML, the authors propose fine grained access control policies, which are used to specify the instances of a concept that a user is permitted to access. The proposed fine grained access control policies are enforced over RDF data, by expanding a SeRQL query (an alternative to SPARQL), to include triple patterns for the instances that are permitted.

Ferrini and Bertino [28] describe an extension of XACML, which uses a combination of XACML and OWL, in order to support RBAC constraints, such as static and dynamic separation of duty. Like Franzoni et al. [32], XACML policies are specified using OWL, therefore it is possible to take advantage of OWL’s out of the box reasoning capabilities. The authors propose a two layer framework where access control policies are specified using XACML and constraints and role hierarchies are represented using OWL. They further extend the XACML enforcement framework which enhances the XACML engine with Description Logic reasoning capabilities.

2.2.2. Web Identity and Discovery
Web Identity and Discovery (WebID) is supported by a W3C community group, is a mechanism used to uniquely identify and authenticate a person, company, organisation or other entity, by means of a Uniform Resource Identifier (URI) [80]. Essentially a WebID is a HTTP URI which is used to represent an
agent. A description of the agent is provided in an RDF document, known as a WebID profile, which can be dereferenced using 303 redirects or Hash URI’s. The WebID-TLS protocol (where TLS stands for Transport Layer Security) specifies how together a WebID profile and a public key certificate can be used to authenticate users [88]. The user places their WebID profile document URI in the Subject Alternative Names field of their certificate. Once the certificate has been generated the user adds the public key details to their WebID profile document. A service wishing to authenticate the user, needs to verify that the public key of the certificate it receives matches the public key specified in the WebID profile.
Adoption. Hollenbach et al. [39] use FOAF[1] and the Secure Sockets Layer (SSL) to determine if the public key in the users FOAF profile matches that of the certificate. When a requester attempts to authenticate, the system extracts the public key from the certificate. The system subsequently verifies the signature and if successful, a SPARQL query is executed against the FOAF profile, in order to verify that it contains a matching public key.

In Berners-Lee et al. [5], the authors describe their vision of a read-write web of data and present a proof of concept. Like Hollenbach et al. [39], the authors discuss how WebID together with FOAF+SSL, can be used for authentication. It is worth noting that although Sternsek et al. [85] do not use the term WebID, they describe how attributes pertaining to a subject (commonly known as credentials), can be associated with public keys and attached to digital certificates.

2.2.3. Web Access Control

WebAccessControl (WAC)[2] is an RDF vocabulary and an access control framework, which demonstrates how together WebID and access control policies specified using the WAC vocabulary, can be used to enforce distributed access control. WAC authorisations grant agents, access to resources. Agents are specified using the agent and agentClass properties and resources are specified using the accessTo and accessToClass properties. Whereas, Read, Write, Append and Control access rights are represented as classes. Once the user has been authenticated using WebID, the system checks if a policy exists which grants the user access to the requested resource. If no such policy exists, then the system checks for classes that are granted access. For each class the system dereferences the URI and checks if the users WebID is a type of the given class. If yes, then the user is granted access to the system. When it comes to WAC, primary research efforts focus on: (i) demonstrating how WebID and WAC can be used for authentication and authorisation respectively [39]; and (ii) extending the WAC vocabulary to cater for more expressive access control policies [100,75].

Adoption. In addition to using WebID for authentication, Hollenbach et al. [39] use the WAC vocabulary to specify access control policies. The authors provide a mapping between permissions and HTTP operations and demonstrate how together WebID and WAC can be used to grant/deny access to web resources. In order to verify if the user has the permissions required to perform the requested HTTP operation, a SPARQL query is executed against the access control policy. If access is denied, a 403 response is returned from the server.

Extensions. Both Villata et al. [100] and Sacco and Passant [75] extend the WAC to cater for access control over the RDF data model. Using the extended vocabularies, it is possible to associate access control with individual RDF resources (subjects, predicates and objects) and also collections of RDF resources (named graph). In addition, the authors extend the vocabulary to cater for a broader set of access privileges (create, read, write, update, delete, append and control).

2.2.4. Platform for Privacy Preferences

Platform for Privacy Preferences (P3P)[3] is a W3C recommendation, which enables websites to express their privacy preferences in a machine readable format. Like XACML, the specification provides an XML Schema, which can be used to specify policies. In addition, the specification details how privacy policies can be associated with webpages/websites and describes how P3P policies can be used in conjunction with HTTP. Organisations wishing to specify machine readable privacy policies can publish their privacy policies using the P3P syntax. A reference to the policy can be added to a well known location (for example, /w3c/p3p.xml), which can be specified using the HTML link tag, or alternatively can form part of the HTTP Response. P3P agents can be built into browsers, plug-ins or proxy servers. The agent is responsible for fetching the servers privacy preferences and taking some action. This action can vary from simply displaying a symbol, to comparing the servers privacy preferences to those of the client and taking some form of action. A related specification called A P3P Preference Exchange Language (APPEL) [21] presents a vocabulary, which is used by individuals (as opposed to websites) to express their privacy preferences. When it comes to P3P and RDF, primary research efforts focus on: (i) representing P3P policies using OWL Garcia and de Toledo [34]; and (ii) extend-
ing the P3P vocabulary to cater for more expressive privacy preferences Kolari et al. [55].

Adoption. Garcia and de Toledo [34] demonstrate how P3P policies can be represented in OWL. The authors detail how the Web Services Policy Framework (WS-Policy) can be used by service providers and consumers to specify their privacy policies using OWL ontology concepts based on the P3P vocabulary. Policies are associated with web services and access is determined by a broker by comparing consumer privacy preferences with provider privacy guarantees, taking into account the semantic relationships between ontology concepts.

Extension. Kolari et al. [55] propose an extension to P3P, to cater for more expressive privacy preferences. The P3P policies are specified using a policy language based on Semantic Web technologies, known as Rei. As Rei is a general policy language, it can easily be used to represent existing P3P policies in a manner which supports reasoning based on context. As access rights in Rei are based on deontic logic, it is possible to model, not only positive and negative permissions, but also positive and negative obligations.

2.2.5. The Open Digital Rights Language

The Open Digital Rights Language (ODRL) [11] is a general rights language, which is used to define rights to or to limit access to digital resources. The information model and the associated semantics is documented in the core model. Whereas, the common vocabulary provides a dictionary of terms that can be used to express permissions, prohibitions, constraints, and duties with respect to digital assets. In addition, a number of serialisations of the core model are provided e.g. XML, JSON and RDF. The core model may be amended/extended in order cater for disparate policy requirements (i.e. different communities may define their own ODRL profiles). When it comes to ODRL and the RDF specifically, primary research efforts to date focus on: (i) using ODRL vocabularies to specify RDF licenses Cabrio et al. [13]; and (ii) demonstrating how ODRL can be used to express a variety of access policies Steyskal and Polleres [86].

Adoption. Cabrio et al. [13] demonstrate how together the Creative Commons Rights Expression Language Ontology can be used to specify Creative Commons (CC) licenses and how the ODRL Ontology can be used to model all other licenses. In addition, the authors use Natural Language Processing (NLP) techniques to automatically generate machine readable licenses (represented using RDF) from their natural language counterparts. Steyskal and Polleres [86] examine the suitability of ODRL for specifying access policies for Linked Data. The authors present several motivating scenarios and demonstrate via example how ODRL can be used to: (i) restrict access to specific datasets; (ii) limit the number of permissible requests; (iii) grant or deny access for a specific time window; (iv) represent common licenses; (v) limit data reuse; and (vi) define policies that require the payment of duties.

3. General Policy Languages and Frameworks

Policy languages can be categorised as either general or specific. In the former, the syntax caters for a diverse range of functional requirements (access control, query answering, service discovery, negotiation, to name but a few), whereas the latter focuses on just one functional requirement. Two of the most well-known access control languages, KAoS [11, 29] and Rei [47, 48], are in fact general policy languages. Natural language, programming languages, XML and ontologies can all be used to express policies. XML and ontologies are two popular choices for representing policy languages as they benefit from flexibility, extensibility and runtime adaptability. However, ontologies are better suited to modelling the semantic relationships between entities. Also, the common framework and vocabulary used by ontologies, to represent data structures and schemas, provides greater interpretability and interoperability. Regardless of the language chosen, a logic based underlying formalisation is crucial for automatic reasoning over access control policies.

The work presented in this section is limited to policy languages that use ontologies, rules or a combination of both to represent general policies. A summary of the existing proposals is presented in Table 3. While, a detailed timeline is presented in Section 5.

As the objective is to provide the reader with an overview of each approach, a detailed description of well known frameworks in each category is presented. For a broader comparison of policy languages, the author is referred to a survey by Bonatti and Olmedilla [8]. The authors divide the criteria into two categories:
Table 3
General Policy Languages - Policy Representation and Enforcement

<table>
<thead>
<tr>
<th>Policy Type</th>
<th>Policy Representation</th>
<th>Enforcement Mechanism</th>
<th>Enforcement Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAoS [11,12,95]</td>
<td>± authorisations</td>
<td>DAML &amp; OWL</td>
<td>DL based enforcement, conflict resolution &amp; harmonisation</td>
</tr>
<tr>
<td></td>
<td>± obligations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rei [47,48]</td>
<td>± authorisation</td>
<td>RDFS, Prolog Rules &amp; OWL</td>
<td>rule based enforcement, dynamic constraints, runtime variables, conflict resolution via metapolicies</td>
</tr>
<tr>
<td></td>
<td>± obligations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protune [10,7,8]</td>
<td>decision, provisional &amp; abbreviation predicates</td>
<td>lightweight ontologies, rules and meta rules</td>
<td>rule based enforcement, disclosure &amp; negotiation</td>
</tr>
<tr>
<td>Proteus [92]</td>
<td>-</td>
<td>policies and domain info as classes, user context as instances</td>
<td>DL &amp; rule based enforcement, conflict resolution &amp; harmonisation, dynamic constraints, runtime variables, disclosure &amp; negotiation</td>
</tr>
<tr>
<td>Kolovski et al. [56]</td>
<td>-</td>
<td>XACML policies as DL</td>
<td>disclosure, rules for conflict resolution</td>
</tr>
</tbody>
</table>

the more theoretical criteria, which they refer to as core policy properties (well-defined semantics, monotonicity, condition expressiveness, underlying formalism) and the more practically oriented criteria, which they refer to as contextual properties (action execution, delegation, type of evaluation, evidences, negotiation support, policy engine decision, extensibility). However the criteria was devised in order to compare existing policy languages that focus on access control and trust, therefore it would need to be amended depending on the access control requirements of the specific use cases. In Section 5 we present an extended list of requirements that are derived from several papers that examine access control policies, with standard description logic reasoners. This section examines KAoS [11,12,95] a general policy language which adopts a pure ontological approach.

3.1. KAoS

KAoS [11,12,95] is an open distributed architecture, which allows for the specification, management and enforcement of a variety of policies. In initial versions of the language, policies were represented using DAML [13]. However, the authors later moved to OWL, the successor of DAML [99]. As both DAML and OWL are based on description logic, using the KAoS language it is possible to define class and property hierarchies, along with inference rules. Although KAoS was originally designed to enable interoperability between complex web agents (software that acts on behalf of humans) [11,89], it was later applied to web services [96,97,99,98] and grid computing [45,96].

Policy Specification. The authors define a set of core vocabularies, known as KAoS policy ontologies, that is used to specify policies. A policy is used to express either an authorisation or an obligation, on the part of one or more actors, with respect to actions relating to resources. Policies are represented as instances of the aforementioned policy types. The language is not meant to be exhaustive, but rather to provide a basis, which can be further extended, to cater for use case specific classes, instances and rules. In order to simplify policy administration and enforcement, actors and resources are organised into domains, that can be nested indefinitely. Domains and subdomains are used to represent complex relationships between classes and

---


instances, such as organisation structures.

**Enforcement of policies.** The authors propose a general policy and domain services framework, which consists of the following components: a policy administration tool, directory services, guards, enforcers and a domain manager. The policy administration tool, known as KPAT, is a user friendly interface that allows administrators, who are unfamiliar with DAML and OWL, to either specify new or maintain existing policies. Guards are responsible for enforcing platform independent policies. While, enforcers are responsible for enforcing policies that are platform dependent. The domain manager is used to manage domain membership and to distribute policies to guards. This component is also responsible for notifying the guards of any policy updates. Given that the actual enforcement may depend not only on the action to be performed, but also on the application, it may be necessary for the developer to implement platform specific code. In order to simplify the integration of these custom enforcement mechanisms with the KAoS framework a number of interfaces are provided as a guide for developers.

In a follow up paper [94], the authors discuss how description logic can be used to support policy administration, exploration and disclosure. From an administration perspective the authors are primarily concerned with subsumption based reasoning and the determination of disjointness. The former is used to investigate if one class is a more general form of another (e.g. in role hierarchies an IT manager is also an employee) and the latter is used to ensure that sets of subjects, access rights and objects are mutually exclusive (e.g. a user cannot belong to an administrative roles and a non-administrative role). Using abductive reasoning it is possible to both test constraints, and to return relevant constraints given one or more properties. Using deductive reasoning it is possible to identify and resolve conflicts at design time. The authors propose a general algorithm for conflict resolution and harmonisation, which can be used even when the entities (actors, actions, resources and policies) are specified at different levels of abstraction. The proposed conflict resolution strategy is based on policy priorities and timestamps. In the event of a conflict the algorithm takes the policy with the lowest precedence and subdivides it until the conflicting part has been isolated. The conflicting policy is removed, and non conflicting policies are generated and feed into the knowledge base.

### 3.2. Rule Based Approaches

One of the benefits of a rule based approach is that it is possible to support access control policies that are dependent on dynamic constraints that can only be evaluated at run time. Like ontology based approaches, access control policies are defined over ontology entities. This section examines two different rule based languages and enforcement frameworks, Rei [47,48] and Protune [10,78].

#### 3.2.1. Rei

Rei [47,48] is a Semantic Web policy language and distributed enforcement framework, which is used to reason over policies, that are specified using RDFS or Prolog rules. As OWL has a richer semantics than RDFS, the authors later provided an OWL representation for their policy language [24,49]. Like KAoS, Rei is a general policy language which can be applied to agents and web services [24,46]. Although Rei policies can be represented using RDFS or OWL, the authors adopt a rule based enforcement mechanism, in contrast to the description logic enforcement mechanism adopted by KAoS.

**Policy Specification.** Rei provides support for four distinct policy types, permissions, prohibitions, obligations and dispensations. Whereby, permissions and prohibitions in Rei are equivalent to positive and negative authorisations in KAoS, and likewise obligations and dispensations in Rei and equivalent to positive and negative obligations in KAoS. By choosing to represent access rights as speech acts Rei is able to support not only a wide range of policies but also the delegation and revocation of policies. A policy is composed of a set of rules, based on the four policy types, that are used to associate conditions with actions. A has predicate is used to associate permissions and obligations with entities. Like KAoS, the core ontologies can be further extended to meet the requirements of specific use cases.

**Enforcement of policies.** The Rei policy framework, called Rein (Rei and N3) [46,49], consists of the following components:

- a set of ontologies, used to represent the Rein policy network, which is composed of the Rein policy language, resources, policies and metapolicies (additional rules over the policy language commonly used to specify defaults and to resolve conflicts) and access requests; and
– a reasoning engine, that uses both explicit and derived knowledge to determine if a request should be granted or denied.

The authors propose a distributed enforcement architecture, whereby each entity is responsible for specifying and enforcing their own policies. Rein is capable of acting as a server or a client. In server mode, Rein retrieves the relevant policies; requests the credentials necessary to access the resource; and verifies the credentials against the policies. Whereas in client mode, the server returns a link to a policy which the client must satisfy; the Rein client generates a proof that the requester can satisfy the policy; and forwards the proof to the server. In order to cater for scenarios where part of the policy is private and part is public, the authors propose a hybrid approach, where Rein acts both as a client and a server. For example, if a company offers a discount for a product/service, the policy will request proof of a valid discount (i.e. the public part), however the details of the relevant discounts will not be disclosed (i.e. the private part). The details of how such a hybrid approach would work in practice are left to future work.

Using Rein it is possible to combine and reason over different access control policies, metapolicies and policy languages. Policies are expressed using either RDFS or OWL, and inference over both data resources and policies is performed using an N3 reasoner, known as Cwm. N3 was originally used as a representation syntax for RDF, however it was later extended to allow for variables and nested graphs. Cwm extends N3 with inference rules and built-in functions, making it possible to express relationships between graphs, specify both existential and universal constraints and to represent implication. Although the authors demonstrate how the Rei vocabulary can be used to specify policies, these policies could be represented using alternative vocabularies.

In Kagal et al. [48], the authors discuss how conflict resolution can be achieved using metapolicies. Priority policies are used to indicate dominance between policies or policy rules. While, precedence policies are used to specify a default grant or deny, for policies, sets of actions or sets of entities satisfying specific conditions. In order to guarantee that a decision can always be reached, the authors propose a partial order between metapolicies. Given Rei allows for policies to contain variables, conflicts need to be resolved at run-time, as opposed to design-time, in the case of KAoS.

3.2.2. Protune

Protune [10,78] is a policy language which was proposed by the Research Network of Excellence on Reasoning on the Web, known as REWERSE. Like Rei, Protune adopts a rule based approach to policy enforcement. The authors identify usability as one of the primary factors for a policy aware web. To this end, Protune was designed to support both trust negotiation and policy explanations. Lightweight ontologies are used to represent concepts, the relationships between these concepts and details of the evidences needed to prove their truth. Protune is an extension of two other well known policy languages, the Portfolio and Service Protection Language (PSPL) [9] and PeerTrust [35]. PSPL is a model and framework, which uses rules to support policy filtering, policy exchange and information disclosure. Whereas, PeerTrust is a language and a framework, which uses semantic annotations and access control rules, in order to cater for automated trust negotiation and access control.

Policy Specification. Protune policies are specified using rules and meta-rules (essentially horn clauses with some syntactic sugar), which provide support for both deductive and abductive reasoning. The former is used to enforce policies, whereas the latter is used to retrieve information about the policy conditions that need to be satisfied. Protune provides three predicate categories (decision predicates, provisional predicates and abbreviation predicates).

– Decision predicates are used to specify the outcome of a policy.
– Provisional predicates are used to represent the conditions the requester must satisfy. By default the system supports two conditions: requests for credentials and request for declarations. Both credentials and declarations are used to assert facts about the requester, however credentials are certified by a third party, whereas declarations are not.
– Abbreviation predicates, which are composed of one or more provisional predicates, are used to represent abstractions of the conditions listed in the body of the rule, simplifying policy specification and maintenance.

15REWERSE, http://rewerse.net/
It is also possible to extend the language with custom predicate categories. Ontologies are used to associate evidences (descriptive requirements of what is needed to meet the conditions) with access conditions. Evidences of this nature facilitate negotiation.

In protune metarules are used to specify constraints and to drive negotiation decisions. For instance using metarules it is possible to attach a provisional predicate to an action in order to specify the actor who is permitted to execute the action, or to assign sensitivity levels to predicates and rules that can be used for policy filtering (i.e. to remove the sensitive parts of the policy).

**Enforcement of policies.** The enforcement framework is composed of three separate components, a *negotiation handler*, an *execution handler* and an *inference engine*.

- The *negotiation handler* is responsible for sending conditions to the requester and providing responses to conditions that were requested.
- The *execution handler* is used to interact with external systems and data sources.
- The *inference engine* is tasked with both enforcing policies (deduction) and retrieving evidences (abduction).

Like Rei, Protune can be used as a client, as a server, or both. Protune-x is a key component of the Protune framework, which provides policy explanations in controlled natural language. Using verbalization metarules it is possible to specify in controlled natural language how domain-specific atoms have to be rendered. For example using the following rule it is possible to explain that Y is the password of X:

\[
\text{passwd(X, Y) } \rightarrow \text{verbalization : Y &” is the password of” &X.}
\]

Protune-x supports four different types of queries:

- *How-to queries* (provide a description of the policy).
- *What-if queries* (give foresight into potential policy outcomes).
- *Why queries* (give explanations for positive outcomes).
- *Why-not queries* (give explanations for negative outcomes).

Protune is developed in Java with a Prolog reasoning component, which is compiled into Java byte code.

### 3.3. Hybrid Approaches

A hybrid approach to policy specification and enforcement can be used to exploit the out of the box deductive capabilities, of an ontology based approach, and the runtime inference capabilities, of a rule based approach. This section describes Proteus [92] which uses a combined approach to policy enforcement and examines an alternative approach, presented by Kolovski et al. [56] which demonstrates how description logic based access control policies can be extended with defeasible logic rules.

#### 3.3.1. Proteus

Proteus [92] uses a hybrid approach to access control policy specification. The authors examine early versions of KAoS and Rei, and highlight the strengths and weaknesses of both ontology based and logic based policy languages and frameworks. Like KAoS the authors use ontologies to model both domain information and policies. Such an approach allows for conflict resolution and harmonisation at design time. Like Rei, the authors adopt a rule based approach in order to support dynamic constraints and run time variables. For example, to support access control based on dynamic context pertaining to the requester or the environment. Like Protune, policy descriptions are used to facilitate partial policy disclosure and policy negotiation.

**Policy Specification.** Policies are represented as classes and contextual information, relating to the user, are represented as instances. Description logic deduction is used to determine the policies that are relevant for the instance data supplied. However, using description logic reasoning it is not possible to cater for contextual properties that are based on property paths or that are associated with variables. In order to handle reasoning of this nature, the authors propose context aggregation and context instantiation rules. Such rules are represented as horn clauses, with predicates in the head and ontological classes and properties in the body.

**Enforcement of policies.** The Proteus policy framework [93] is composed of the following core components: a *policy installation manager*, a *reasoning core*, a *policy enforcement manager* and a *context manager*.

- The *policy installation manager* is responsible for loading ontologies, access control policies, contextual information and quality constraints.
– The reasoning core performs reasoning over policies, context and quality constraints in order to determine which policies are currently active.
– The policy enforcement manager intercepts action requests, collects relevant contextual information and interacts with the reasoning core in order to determine if access should be granted or denied.
– The context manager collects state information pertaining to system entities and forwards this contextual information to the reasoning core.

The authors provide details of their prototype which is implemented in Java with a Pellet reasoner. The proposed solution supports incremental reasoning via an OWL application programming interface and SPARQL queries.

3.3.2. Kolovski et al. [56]

Kolovski et al. [56] demonstrate how together description logic and defeasible logic rules, known as defeasible description logic [36], can be used to understand the effect and the consequence of sets of XACML access control policies.

Policy Specification. The XACML policies are represented using description logic and the combination algorithm, which is used to handle conflicts is represented using defeasible description logic [36]. The formalism supports both strict rules that cannot be overridden and defeasible rules that may be overridden by a higher priority rule.

Enforcement of policies. Although the actual enforcement framework is not presented, the following subset of policy services are described:
– Constraints. Like Finin et al. [29,30], the proposed solution caters for role cardinality and separation of duty;
– Comparison. Policies or sets of policies can be compared in order to determine if one is equivalent to or logically contains the other;
– Verification. Like Bonatti and Olmedilla [8], this component checks if the policy satisfies a given property;
– Incompatibility. This component provides details of policies that cannot be active at the same time;
– Redundancy. This component checks hierarchies to ensure that all policies are reachable; and

– Querying. Given a set of attributes, this component searches for relevant policies.

The proposed XACML analysis prototype is implemented on top of Pellet (an open source description logic reasoner).

4. Access Control for RDF

This section presents the different access control mechanisms that have been used to protect RDF data. In particular, it focuses on the specification of access control policies, the different enforcement mechanisms, the simplification of administration using different reasoning strategies and the alternative techniques used to return partial query results. A summary of the existing proposals is presented in Table 4. While, a detailed timeline is presented in Section 5.

4.1. Specification of Access Control for RDF

Over the years several researchers have focused on specification of access control policies over RDF data. A number of authors [70,41,1,31] define access control policies based on RDF patterns, that are mapped to one or more RDF triples. Others [57,33,17] propose view based access control strategies for distributed RDF data. Whereas, [75,76,77,74,100,17,18,86,87] propose access control ontologies.

4.1.1. RDF Patterns

Triple patterns are triples that can potentially contain variables in the subject, predicate and object positions. From an access control perspective triple patterns are used to match multiple triples.

In Reddivari et al. [70] two predicates permit and prohibit are used to grant and deny access rights to one or more triples using triple patterns. Access control policies are defined using Prolog facts and rules and compiled into Jena rules. Authorisations can be further constrained using conditions relating to policies, triples and agents:

– policy specific conditions relate to the access control policies, for example a user can only add instances if they added the class.
– triple specific conditions correspond to the triple specified in the authorisation, for example if an authorisation governs a triple then all triples associated with a subProperty relation are governed by the same policy.
Abel et al. [1], Flouris et al. [31] and Kirrane et al. [53,52] also use RDF triple patterns to expose or hide information represented as RDF. However, Abel et al. [1] and Flouris et al. [31] go beyond simple graph patterns by allowing the graph pattern to be constrained by a \texttt{WHERE} clause.

Jain and Farkas [41] also use triple patterns to specify access control policies, however the authors build on the approach proposed by Reddivari et al. [70], by demonstrating how RDFS entailment rules can be used to derive access rights for inferred triples (see Section 4.4).

4.1.2. Views and Named Graphs

An RDF graph is a finite set of RDF triples. Named graphs are used to collectively refer to a number of RDF statements. A collection of RDF graphs, which can include a default graph and one or more named graphs is known as an RDF dataset.

Gabillon and Letouzey [33] highlight the possible administration burden associated with maintaining access control policies that are based on triple patterns. They propose the logical distribution of RDF data into views using SPARQL \texttt{CONSTRUCT} and \texttt{DESCRIBE} queries and the subsequent specification of access control policies, based on existing RDF graphs or predefined views. Access control policies are specified using contextual information pertaining to the user, resources or the environment. The body of the rule is a possibly empty condition, or a combination of conditions connected via conjunction or disjunction. The head of the rule is an authorisation. Like Gabillon and Letouzey [33], Dietzold and Auer [25] describe how RDF data can be logically organised into views using SPARQL \texttt{CONSTRUCT} queries.

4.1.3. Ontology Concepts

Sacco and Passant [75,76] and Sacco et al. [77] demonstrated how an extension of the Web Access Control vocabulary known as the Privacy Preferences Ontology (PPO) can be used to restrict access to an RDF resources, statements and graphs. An access control policy is composed of:

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|}
\hline
Policy Representation & Enforcement Framework & \\
\hline
Abel et al. [1] & RDF patterns with \texttt{WHERE} clause & query rewriting via bindings \\
\hline
Amii and Jalili [1] & - & reasoning framework \\
\hline
Ehsan et al. [32] & - & safe reasoning \\
\hline
Bao et al. [4] & ontology triples, classes and properties & using SPARQL views for data filtering \\
\hline
Chen and Stuckenschmidt [26] & - & query rewriting via \texttt{FILTERS} \\
\hline
Dietzold and Auer [25] & RDF patterns & flexible authorisation framework \\
\hline
Flouris et al. [31] & RDF patterns with \texttt{WHERE} clause & - \\
\hline
Franconi et al. [52] & - & query rewriting via bindings \\
\hline
Gabillon and Letouzey [33] & RDF patterns & SPARQL views for data filtering \\
\hline
Jain and Farkas [41] & - & RDF inference to deduce authorisations \\
\hline
Javanmardi et al. [43] & RDF patterns & - \\
\hline
Kim et al. [59] & RDF patterns & - \\
\hline
Kirrane et al. [53,52] & - & reasoning over ontology concepts, properties and individuals \\
\hline
Li and Cheung [57] & views & query rewriting via expanded views \\
\hline
Lopes et al. [58] & RDF patterns & - \\
\hline
Muhleisen et al. [50] & ontology triple patterns, resources and instances & RDFS inference and rights propagation \\
\hline
Oudmakhche et al. [63] & - & query rewriting via \texttt{FILTERS} \\
\hline
Papakonstantinou et al. [63] & RDF patterns & RDFS inference to deduce annotations \\
\hline
Qin and Adluri [69] & - & reasoning over ontology concepts \\
\hline
Reddivari et al. [70] & RDF patterns & - \\
\hline
Rykov et al. [72] & using ontologies to extend WAC & SPARQL \texttt{ASK} queries \\
\hline
Sacco et al. [77] & using ontologies to extend WAC & SPARQL \texttt{ASK} queries, \\
\hline
Sacco and Passant [75,76] & using ontologies to extend WAC & query rewriting using named graphs \\
\hline
Steyskal and Felli [66,67] & using ontologies to represent ODRL & \\
\hline
Villata et al. [78] & using ontologies to extend WAC & \\
\hline
Costabello et al. [79] & using ontologies to extend WAC & \\
\hline
\end{tabular}
\caption{Access Control for RDF - Policy Representation and Enforcement}
\end{table}
– a restriction in the form of an RDF resource, statement or graph;
– a condition which provides specific details of the restriction, for example hasProperty, hasLiteral;
– an access privilege, for example read and/or write access rights; and
– a SPARQL ASK query that must be satisfied by the requester.

A follow-up paper Sacco and Breslin \cite{Sacco2009} extends the original PPO to allow access to be restricted based on a dataset or a particular context. The authors also provide support for more expressive authorisations (in the form of negation and logical operators), and a broader set of access privileges (create, read, write, update, delete, append and control).

An alternative access control vocabulary called Social Semantic SPARQL Security for Access Control (S4AC) is presented in Villata et al. \cite{Villata2016}. Like Sacco and Passant \cite{Sacco2008} they extend the WAC to cater for fine grained access control over RDF data. Their proposal is tightly integrated with several social web and web of data vocabularies. The authors define access control policies for named graphs, which can also be used to grant/deny access to sets of triples. S4AC provides support for logical operators and a broad set of access privileges (create, read, write, update, delete, append and control) from the offset.

In contrast Steyskal and Polleres \cite{Steyskal2012} discuss how an alternative vocabulary known as the Open Digital Rights Language (ODRL) 2.0 ontology can be used to specify access policies for Linked Data. The ODRL vocabulary is more general than the WAC vocabulary, and thus in addition to standard access control policies the authors demonstrate how ODRL can be used to represent common licenses and policies that require payment. In a follow-up paper \cite{Steyskal2015}, the authors propose a formal semantics for ODRL, which can be used as the basis for rule based reasoning over ODRL policies.

4.2. Enforcement of Access Control for RDF

Over the years several researchers have focused on the modelling and enforcement of access control over RDF data. Existing proposals can be summarised as follows: (i) demonstrating how access control can be enforced on top of RDF Data \cite{Reddivari2006, Flouris2007}; and (ii) enforcing access control over access control policies represented as RDF using SPARQL ASK queries \cite{Sacco2009, Villata2016}.

4.2.1. Policy Layer

Reddivari et al. \cite{Reddivari2006} define a set of actions required to manage an RDF store and demonstrate how access control rules can be used to permit or prohibit the requested actions. The actions are organised into four categories:

– adding (the insertion of explicit triples, implicit triples and sets of triples);
– deleting (the deletion of explicit triples, implicit triples and sets of triples);
– updating (directly replacing one triple with another); and
– querying (returning triples or using triples to return answers to queries).

The proposed RDF Store Access Control Policies (RAP) framework checks the policy to ensure that the action is permitted, temporarily allows the action, and afterwards checks the policy to ensure that the inferences are allowed. The authors propose default and conflict preferences that can simply be set to either permit or deny.

Like Reddivari et al. \cite{Reddivari2006}, Flouris et al. \cite{Flouris2007} propose a default policy and a conflict resolution strategy. They formally define the semantics of the individual access control statements and the entire access control policy, and present the different possible interpretations for the default semantics and the conflict resolution. A flexible system architecture that demonstrates how the access control enforcement framework can be used with disparate RDF repositories and query languages is presented. The system is implemented using Jena ARQ, Jena SDB with a Postgresql back-end and Sesame.

Gabillon and Letouzey \cite{Gabillon2010} describe an enforcement framework, whereby users define security policies for the RDF graph/views that they own. Users may delegate rights to other users by specifying an authorisation which grants construct and describe privileges to the RDF graph or view. Although the authors acknowledge the need for conflict resolution, they do not propose a conflict resolution strategy.

4.2.2. SPARQL ASK queries

Sacco and Passant \cite{Sacco2008, Sacco2009} and Sacco et al. \cite{Sacco2010} describe the formal semantics of the PPO and present a detailed description of their Privacy Preferences Manager (PPM), which can be used to enforce access control using SPARQL ASK queries. A follow-up paper Sacco and Breslin \cite{Sacco2009} extends the original PPO and PPM to allow access to be restricted based on
a dataset or a particular context. The authors also provide support for conflict resolution, more expressive authorisations (in the form of negation and logical operators), and a broader set of access privileges (create, read, write, update, delete, append and control). In Sacco et al. [78], the authors demonstrate how both the PPO and the PPM can be used to cater for fine-grained access control on a mobile device. A number of shortcomings of their original enforcement algorithm are identified and a more efficient algorithm which utilises pre-indexing, query analysis and results filtering is presented and evaluated.

Like Sacco and Passant [75], Villata et al. [100] use SPARQL ASK queries to determine if the requester has the permissions necessary to access a resource. The authors propose the disjunctive evaluation of policies, thus circumventing the need for a conflict resolution mechanism. Follow up work by Costabello et al. [17] describes how an access control framework, called Shi3ld, can be used to enforce access control over SPARQL endpoints in a pluggable manner. In [19] the authors extend the Shi3ld framework to cater for access control for a Linked Data Platform (LDP). Resources refer simply to Linked Data resources that are queries, created, modified and deleted via HTTP requests processed by a LDP. Two alternative frameworks are presented, one which contains an embedded SPARQL engine and a SPARQL-less solution. In the first scenario, Shi3ld remains unchanged. Whereas in the second scenario, authorisations cannot contain embedded SPARQL queries and therefore are evaluated using subgraph matching. In their evaluation the authors compare all three frameworks, using the Billion Triple Challenge 2012 Dataset. Based on their performance evaluation the authors conclude that access control over SPARQL endpoints is marginally slower than access control over LDP resources and that their SPARQL-less solution exhibits a 25% faster response time.

4.3. Reasoning over RDF Access Control Policies

Inference is a process whereby new data is derived from data which is known or assumed to be true. Section 3 discussed how deduction and abduction can be used to simplify both policy specification and maintenance. However, inference can also be used to deduce information, which users should not have access to, commonly known as the inference problem. Both Thuraisingham [91] and Nematzadeh and Pournajaf [62] highlight the need for security mechanisms to protect against such unauthorised inference. Although this is not a new problem, the authors argue that with advances in current data integration and mining technologies, the problem is further magnified. According to Qin and Atluri [69], if the semantic relationship between entities is not taken into account it may be possible to infer information which has been restricted, or access control policies may not exist for the inferred information making this information inaccessible.

Existing reasoning proposals focus on: (i) demonstrating how access rights can be inferred for new triples deduced based on RDFS inference rules [41,50,65,58]; (ii) the propagation of access rights based on authorisation subjects, access rights and resources [69,43,73,3]; (iii) proposing a flexible authorisation framework [53,52]; and (iv) examining use cases where it is desirable to grant access to data which has been inferred from unauthorised data [4].

4.3.1. RDFS Inference

When it comes to RDFS inference, there are two different strands of research. The first infers access rights for triples that are inferred using RDFS entailment rules [41,50]. Whereas the second uses RDFS entailment rules to propagate permissions for triples that already exist [65,58].

Inferring access rights for new triples. Jain and Farkas [41], demonstrate how RDFS entailment rules can be used not only to infer new RDF triples, but also to infer access control annotations for those triples. The authors use RDF triple patterns and associated security classifications, known as security labels, to limit access to RDF statements. They define a subsumption relationship between patterns and stipulate that subsuming patterns must be as restrictive as the subsumed patterns. In addition, they define a partial order between security labels, which is used to determine the security classification of triples inferred via RDFS entailment rules. If more than one pattern maps to a statement the most restrictive or the lowest upper bound takes precedence. The authors provide formal definitions for each of the RDF security objects and define an algorithm to generate security labels for both explicit and inferred triples based on a security policy.
Lopes et al. [58] demonstrate how AnQL, an extension of the SPARQL query language, can be used to enforce access control, by rewriting using the requesters credentials to rewrite a SPARQL query to an AnQL query. A follow-up paper [51] demonstrates how custom rules can be used to support multiple access control models and demonstrate how rules can be used to propagate permissions: based on hierarchies of authorisation subjects, access rights and resources; to triples with the same subject; and using resource typing.

Papakonstantinou et al. [65] evaluate their prototype over both ProgressSQL and MonetDB relational databases. Based on their performance evaluation of both the inference and propagation rules, the authors concluded that more efficient storage and indexing schemes are required.

4.3.2. Propagation of Authorisations

Early proposals for the propagation of authorisations focused on reasoning over ontology concepts [69]. Subsequent work by Javanmardi et al. [44,43] focused not only on ontology concepts but also on reasoning over ontology properties and ontology instances. An alternative strategy which is proposed by Ryutov et al. [72,73], takes a more abstract approach by propagating policies based on nodes and edges in a semantic network.

Ontological reasoning. Qin and Atluri [69] extend XML based access control to take into account the semantic relationships between the concepts that need to be protected. The authors propose a Concept Level Access Control (CLAC) model, which allows for reasoning over concepts appearing in authorisation subjects, permissions and objects. Access control policies are represented using an OWL based vocabulary, which they call the Semantic Access Control Language (SACL) and data instances are defined in domain ontologies. The authors use two properties SACL:higherLevelThan and SACL:lowerLevelThan to specify a partial order between authorisation subjects and permissions. The proposed access control propagation is based on six domain independent relationships (i.e. superclass/subclass, equivalence, partof, intersection, union, complement). In the case of equivalence, partof, union and subclass positive policies are propagated from subject to object and negative policies are propagated from object to subject. Where there is an intersection between two concepts, only negative policies are prop-

Inferring and propagating access rights. Both Lopes et al. [58] and Papakonstantinou et al. [65] propose strategies for deriving access control annotations using RDFS inference policies. In the respective access control models both the triples and the corresponding annotations are represented as quads. Annotations can be:

- directly associated with triples;
- inferred using RDF inference rules;
- propagated using RDF inference rules; or
- assigned a default label.

The authors demonstrate how the RDFS subClass, subProperty and type inference rules can be used to assign annotations to inferred triples. However, they do not dictate how the access control annotations assigned to the premises should be combined, but rather propose an abstract operator which can be adapted to suit particular use cases. In addition, the authors demonstrate how the RDFS subClass, subProperty and type inference rules can be used to propagate permissions to existing triples. As per inference rules existing annotations and propagated annotations are inferred by means of a domain operator.

and a conflict resolution strategy. Limited details of the implementation are supplied and no evaluation is performed.

While, Kim et al. [50] demonstrate how together authorisations and RDFS inference rules can be used to generate new authorisations as opposed to simply access control annotations in the case of Jain and Farkas [41]. An authorisation is defined as a four tuple \( \langle \text{sub, obj, sign, type} \rangle \), where \text{sub} refers to the access control subject; \text{obj} is represented an RDF triple pattern; \text{act} is the operation; \text{sign} indicate if access is granted or denied; and \text{type} which is either \text{R} to indicate that the authorisation should be propagated or \text{L} if it should not. The authors discuss how \text{rdfs:subClass}, \text{rdfs:subProperty} and \text{rdf:type} inference can be used to infer new authorisation objects (triple patterns) and consequently new authorisations. In addition, they examine the different scenarios which might result in access to data being both permitted and prohibited. If the authorisation that is prohibited is more specific than the authorisation that is permitted based on the subclass/subproperty hierarchy then access should be denied. In order to determine if there is a conflict, only authorisations with a superclass or superproperty that is negative need to be checked.
agated. In the case of complement relations neither positive nor negative authorisations are propagated. The authors acknowledge the need for conflict resolution and simply propose a negation takes precedence handling mechanism.

The Semantic Based Access Control Model (SBAC) proposed by Javanmardi et al. [44,43], builds on the work presented in Qin and Atluri [69], by catering for access control policy propagation, not only based on the semantic relations between ontology concepts, but also based on the relations between concepts, properties and individuals. Like Qin and Atluri [69], OWL vocabularies are used to represent the authorisation subjects, permissions and objects, however the authorisations themselves are specified using rules. The authors propose the propagation of access rights based on seven different types of inference, from:

- concept to concept (where classes are deemed related based on some vocabulary, for example rdfs:subClass);
- concept to individual (where an entity is a type of class, for example if employee is a class and JoeBloggs rdfs:type employee);
- individual to individual (using properties such as owl:sameAs it is possible to propagate entities that represent the same thing);
- property to concept (if access is granted to the property access should be granted to the classes governed by rdfs:domain and rdfs:range);
- property to property (where properties are deemed related based on some vocabulary. For example rdfs:subProperty or owl:equivalentProperty);
- property to individual (where an entity is a type of property, for example if roles is a property and manager is of rdf:type role); and
- concept to property (where access is granted to a concept it should also be granted to all properties relating to that concept).

The authors describe how the aforementioned semantic relations can be reduced to subsumption relations and propose a general propagation strategy for subsumption relations among subjects, permissions and objects. In the case of subjects and objects, both positive and negative access rights propagate from subsumer to subsumer. However, in the case of permissions positive access rights propagate from subsumer to subsumer, while negative access right propagate from subsumer to subsumer.

Although an architecture is presented by Javanmardi et al. [43], very little detail on the actual enforcement mechanism is supplied. Follow-up papers by Ehsan et al. [26] and Amini and Jalili [3] build on previous work, by providing for an access control model with formal semantics and an enforcement framework, which is suitable for distributed semantic aware environments (for example Semantic Web, Semantic Grid and Semantic Cloud Computing). Policy rules, in both the conceptual and individual levels, are specified using a combination of deontic and description logic, which they refer to as MA(DL). The prototype consists of a user interface developed using the Google Web Toolkit; a data reasoner implemented in Jena; and a tableaux reasoner implemented in Prolog. The authors present the results of a performance evaluation over increasing policy rules, where the decision to grant or deny access is based on ground policies, inferred policies and the proposed conflict resolution strategy. The authors conclude that real-time reasoning is expensive. Therefore they suggest:

- using the parallelisation facilities of the tableaux system;
- adopting a proof based approach where the requester presents authorisation rules that demonstrate they can access the requested resource; and
- materialisation of inferred relations in advance.

Reasoning based on the semantic network. Ryutov et al. [72,73] propose a policy language which can be used to specify access control in terms of the semantic relationships between the nodes and edges of a graph. In order to cater for policy propagation, two directed acyclic graphs are used to represent the relationship between users and groups (using a memberOf property) and between objects and bundles (using a partOf relation). Propagation policies are defined to allow for policy propagation based on both the partOf and memberOf relations. The authors also propose safety and consistency policies that are used to prevent undesirable access control policy specification and propagation, for example resources that nobody can access. The authors propose a conflict resolution algorithm, which is based on the semantic network. When a policy explicitly refers to a node, the policy distance is zero. Whereas, when a policy is implicitly assigned, based on a propagation policy, the distance is determined by counting the number of nodes in the path from the node with the explicit policy. The smaller the distance, the more specific the policy. If multiple policies exist at different distances, the most specific policy takes precedence. If multiple explicit conflict-
ing policies exist, conflicts are resolved using logical conjunction (i.e. if one policy grants access and another policy denies access then the user will be denied access). In the case of implicit policies, conflicts are resolved using logical disjunction (i.e. if one policy grants access and another policy denies access then the user will be granted access).

4.3.3. Flexible Reasoning Framework

Kirrane et al. [53,52] demonstrate how authorisations based on quad patterns together with stratified Datalog rules can be used to enforce DAC over the RDF data model. An RDF quad pattern is an RDF quad with optionally a variable V in the subject, predicate, object and/or graph position. A quad pattern is a flexible mechanism which can be used to grant/restrict access to an RDF quad, a collection of RDF quads (multiple quads that share a common subject), a named graph (arbitrary views of the data), specific classes or properties. The authors describe how the hierarchical Flexible Authorisation Framework proposed by Jajodia et al. [42], which is composed of authorisations, propagation policies, conflict resolution rules and integrity constraints, can be extended to cater for the RDF graph data model. They describe how together pattern matching and propagation rules can be used to ease the maintenance of access control policies for linked data sources; and show how conflict resolution policies and integrity constraints can ensure access control policy integrity. Propagation policies can be used to simplify authorisation administration by allowing for the derivation of implicit authorisations from explicit ones. Rather than propose a conflict resolution strategy the authors provide a formal definition for a conflict resolution rule that can be used to determine access given several different conflict resolution strategies. For example conflict resolution policies based on the structure of the graph data system components; the sensitivity of the data requested; or contextual conditions pertaining to the requester. Integrity constraints are used to restrict authorisation creation based on the existing relationships between SPARQL operations and RDF data items. For example, INSERT and DELETE can only be applied to an RDF quad whereas DROP, CREATE, COPY, MOVE and ADD can only be associated with a named graph. As per conflict resolution the authors provide a formal definition of an integrity constraint and demonstrate how rules can be used to ensure that only valid authorisation and propagation policies can be specified.

4.3.4. Safe Reasoning

When it comes to reasoning over restricted data, there is a general consensus that any information that can be inferred from restricted data should also be restricted. An alternative viewpoint is presented by Bao et al. [4]. The authors focus on a number of use cases where it is desirable to grant access to information that has been inferred from restricted data:

(i) a calendar showing the existence of an appointment without revealing specifics;
(ii) a booking engine sharing partial hotel details; and
(iii) a pharmacy confirming that the patients drugs are reimbursable without disclosing details.

The open world assumption is used to ensure that users cannot distinguish between information which does not exist and information with is inaccessible. The authors stipulate that, the knowledge base should not lie, the answers given should be independent of any previous answers and it should not be possible to infer any restricted data. The authors propose a safe reasoning strategy is based on the notion of conservative extension. Essentially the reasoner keeps a history of the answers to all previous queries. For each subsequent query, the history is consulted, in order to verify that unauthorised information cannot be inferred by the requester.

4.4. Partial Query Results

A number of the access control mechanisms for RDF data that have been presented, demonstrate how their access control can be enforced on top of SPARQL queries. However, the solutions examined thus far either grant or deny access to the entire query. This section examines how data filtering [25,60] and query rewriting [132,14,64,17] can be used to return partial query results when access to some of the results is restricted by an access control policy.

4.4.1. Data Filtering

Dietzold and Auer [25] examine access control requirements for an RDF store from a semantic wiki perspective. The authors propose access control policy specification at multiple levels of granularity (triples, classes and properties). In addition, they define three atomic actions (read, insert and delete) for both individual triples and sets of triples. Authorisations are used to generate a virtual model of the data, upon which user queries are executed. Authorisations are used to associate filters (SPARQL CONSTRUCT queries) with users and resources. When a requester
submits a query, a virtual model is created based on the matched authorisations. The query is executed against the virtual model, which only contains data the requester is authorised to access. Gabrielle and Letouzey [55] also describe how RDF data can be logically organised into views using SPARQL CONSTRUCT and DESCRIBE queries.

Mühlheisen et al. [60] describe a Policy-enabled Linked Data Server (PeLDS), which uses WebID to authenticate users. The policy language caters for the specification of access control policies for particular triple patterns, resources or instances, using SWRL rules. An OWL ontology is used to identify the rule types (single concept and triple pattern) and supported actions query and update. Negation is not supported in the presented modelling. When a requester submits a query, the system uses their WebID to determine the data instances that the user has been granted access to and generates a temporary named graph containing authorised data. The requesters query is subsequently executed against the temporary named graph and the results are returned to the user.

4.4.2. Query Rewriting

Existing query rewriting strategies involve: limiting the query to a specific named graph [17]; rewriting a view so that it considers propagation rules and both instance and range restrictions [57]; and creating bindings for variables and adding them to the query WHERE clause [11, 32, 14, 64].

Named graph added to the query. Costabello et al. [17] restrict access to named graphs using query rewriting. An access control policy is a tuple: \( (\text{ACS}, \text{AP}, S, R, \text{AEC}) \), where \( \text{ACS} \) is a set of access conditions (specified using SPARQL ASK queries); \( \text{AP} \) is a set of access privileges (CREATE, READ, UPDATE or DELETE); \( S \) denotes the subjects to be protected; \( R \) represents the named graphs to be protected; and \( \text{AEC} \) is the evaluation context specified using name value pairs (verified using SPARQL BINDINGS). In addition to the SPARQL query that the user wishes to execute, the user provides their access credentials, in the form of a SPARQL UPDATE query, which contains contextual data. The enforcement framework stores the contextual data in a named graph and retrieves the authorisations that match the query type. In order to determine if access is permitted, the ASK query and the BINDINGS, that are specified in the authorisation, are executed against the users contextual graph. If the ASK query returns true then the query is rewritten to include the corresponding named graph.

Expanding views based on propagation rules and instance and range restrictions. Li and Cheung [57] propose a query rewriting strategy for views generated from ontological relations. An access control policy is defined as a tuple \( (s, v, \text{sign}) \), where \( s \) denotes the subject, \( v \) represents a set of concepts, relations and filters and \( \text{sign} \) is used to indicate permissions and prohibitions. Both views and queries that are also generated from sets of concepts, relations and filters are represented using rules. Propagation policies are used to generate implicit authorisations from explicit authorisations, based on subsumption relations between access control subjects and subsumption relations between concepts, appearing in the body of the view. The proposed query rewriting strategy involves: (i) retrieving the policies applicable to the subject, taking into account the subject propagation rules; (ii) expanding each of the concepts in the body of the view based on the concept propagation policies; and (iii) applying the relevant range and instance restrictions to the query based on the expanded view.

Adding bindings to the query. Abel et al. [11] propose a combined approach to access control enforcement. Contextual conditions that are not dependent on RDF data are evaluated by a policy engine. Whereas the query is expanded to include the contextual conditions that are dependent on RDF data. Such an approach requires the substitution of variables to ensure uniqueness, however in doing so they are able to leverage the highly optimized query evaluation features of the RDF store. In the presented modelling, both positive and negative authorisations are composed of sets of contextual predicates, path expressions and boolean expressions. Queries are assumed to have the following structure SELECT/CONSTRUCT RF FROM PE WHERE BE, where \( RF \) represents the result form (projections in the case of SELECT queries and triples in the case of CONSTRUCT queries); \( PE \) denotes the path expression; and \( BE \) corresponds to one ore more boolean expressions connected via conjunction or disjunction operators. An authorisation is deemed applicable if the triple pattern the policy is protecting, is part of either the \( PE \) or the \( BE \), and the corresponding contextual predicates, path expressions and boolean expressions are satisfied. The authors propose a query rewriting algorithm, which constructs bindings for authorisation path expressions and contextual predicates. For positive authorisations the bindings are appended to the
Franzoni et al. [32] propose a query rewriting strategy, which is used to grant/deny access to ontology instances. The authors rewrite queries to take into account contextual information, pertaining to the user or the environment. A fine-grained access control (FGAC) policy is defined as a tuple: \( \langle \text{target}, \langle \text{property}, \text{attribute}, \text{operator} \rangle \rangle \) where:

- \text{target} is the resource that the policy relates to;
- \text{property} is a path expression, which either directly or indirectly relates to the target;
- \text{attribute} is the user attributes, that are bound to the path expression variables; and
- \text{operator} is the filter condition.

The authors propose a two-tiered approach to access control enforcement. Access control policies are used to determine if access should be granted or denied. FGAC policies are only applied if access is granted. If the query contains one or more FGAC policy targets, the query is rewritten to include the path expression and a WHERE clause, which is composed of an expression generated from the variables in the path expression, the attributes of the requester and the operator.

Chen and Stuckenschmidt [14] present a query rewriting strategy, which can be used to restrict access to data represented using ontologies. The authors focus on restricting access to instance data. Access control policies are used to deny access to specific individuals or to grant/deny access to instances associated with a given class or property. When access is prohibited to specific individuals, a FILTER expression is generated, which ensures that none of the query variables bind to the prohibited individuals. When access is granted to predicates or classes, a FILTER expression is generated, which binds the relevant variables in the query to the specified predicate or class. Whereas, when access is prohibited to predicates or classes, the matching triple patterns are made OPTIONAL and a FILTER expression is generated, which ensures that the corresponding variables do not bind to the specified predicate or class, and variables that are !BOUND are not returned.

Oulmakhzoune et al. [64] propose a query rewriting strategy for SPARQL queries. In the presented modelling, both positive and negative authorisations are composed of sets of filters that are associated with simple conditions or involved conditions. Given a SPARQL query the algorithm examines each individual basic graph pattern (BGP). In the case of \textit{simple conditions}, when authorisations permit/deny access to a single triple pattern, the following query rewriting strategy is applied: If all authorisations that match the triple pattern, permit access to the triple pattern, no action is required; If all authorisations prohibit access to the triple pattern, the pattern is deleted; Otherwise, if the BGP is converted to an OPTIONAL BGP, and the authorisation FILTER expression is added to the query. In the case of \textit{involved conditions}, where authorisations permit/deny access for a given predicate, the following query rewriting strategy is applied: For positive authorisations, if the query contains a triple pattern which matches the predicate of the authorisation, the FILTER condition is added. Alternatively both the triple pattern and the corresponding FILTER are added; For negative authorisations, if the query contains a triple pattern which matches the predicate of the authorisation, and the object is a variable, both the FILTER condition and a !BOUND expression are added; Alternatively the triple pattern, the corresponding FILTER condition and a !BOUND expression are added.

5. Access Control Requirements for Linked Data

More recently, the focus has shifted to the specification and enforcement of access control over Linked Data. Costabello et al. [19], Sacco et al. [77] and Kirrane et al. [52] describe how their policy languages and frameworks can be used in conjunction with Linked Data, and Steyskal and Polleres [86] discuss how ODL can be used to specify access policies for Linked Data. However, any of the access control mechanisms examined thus far could potentially be used (albeit to a lesser or greater extent) to enforce access control over Linked Data.

This section provides a summary of existing requirements for RDF data, and uses these requirements to categorise existing access control strategies that have been proposed for RDF. As the work presented in Section 2 focused on extending existing access control models and standards as opposed to enforcing access control over RDF, the analysis is limited to the policy languages presented in Section 3 and the different access control strategies described in Section 4.
The requirements presented below are derived from several papers that examine access control for RDF from a number of perspectives. Yagüe del Valle et al. [104] examine the different layers of the Semantic Web and how the technologies and concepts can be applied to access control. Both Damiani et al. [22] and Weitzner et al. [102] focus on the access control mechanisms that are required to support new access control paradigms where user privacy is a key requirement. De Coi et al. [23] and Bonatti and Olmedilla [8] investigate the interplay between trust, access control and policy languages. While, Ryutov et al. [73] focus more on the data model, investigating access control requirements from a graph perspective, as opposed to the traditional hierarchical approach. We also consider the set of access control guidelines devised by the W3C Linked Data Platform Working Group (LDP) [90].

Although it would also be interesting to examine the access control requirements arising from emerging Linked Data access strategies (e.g. Linked Data Fragments) and novel languages for navigating and consuming triples on the Web (e.g. nSPARQL), as these technologies are still evolving such an investigation is left to future work.

To ease referenceability the access control requirements are categorised under the headings specification, enforcement, administration and implementation.

5.1. Specification

Generally speaking, access control policy specification requirements relate to the types of policies that can be expressed, and the interoperability of the chosen representation format. An overview of each of the requirements relating to access control specification is presented below and a summary of existing proposals is depicted in Table 5 and the corresponding timeline is represented in Figure 3. One requirement, which was not included is monotonicity. According to Bonatti and Olmedilla [8], the addition of new evidences and policies should not negate any of the previous conclusions. However, given that access control policies may depend on information that can change, for example a change in user context or the removal of a user from a role, it should be possible to update the knowledge base. Therefore, depending on access control model/requirements one could argue that access control should in fact be non-monotonic.

Granularity [23,90]. The W3C Linked Data Platform Working Group [90] identify the need to cater for fine grained access control that supports different levels of granularity (document, named graphs, triples, or individual attributes). Ryutov et al. [73] adopt a graph perspective, stating that it should be possible to specify access control rules for nodes (entities) and edges (semantic relationships between entities). Whereas, Amini and Jalili [3] adopt an ontological view, stating that it is necessary to specify policies for both ontology concepts and individuals. Existing access control strategies for RDF, resources are specified at several different levels of granularity. Namely, triples [25,65,75], named graphs [18,33,75], views [57], triple patterns [41,50,60,70], quad patterns [52], graph patterns with filters [11] and graph patterns without filters [31], classes and properties [25], ontology concepts [4,8,32,43,75], ontology individuals [52,65,75] and graph nodes and edges [73]. In the vast majority of cases, access is either granted or denied. However, a number of researchers have investigated returning partial query results to the requester. Such strategies either involve dataset filtering [25,60] or rewriting the query using either filters [11] or named graphs [18]. In Table 5 the granularity of the authorisations and the different strategies used to cater for partial query results are represented as Granularity and Partial Results respectively.

Underlying Formalism [23,38]. Access control languages should be based on formal semantics, as it decouples the meaning of the policies from the actual implementation. The majority of researchers either adopt formalisms based on logic programming [8,14,93,86] or different flavors of description logic [3,4,93,95]. Kagal et al. [47] demonstrate how logic programming can be combined with deontic logic. Kirrane et al. [52] adopt a DATALOG formalism, Amini and Jalili [3] demonstrate how description logic can be combined with deontic logic, Kolovski et al. [56] combine description logic and defeasible logic and Ryutov et al. [73] adopt a many sorted first order logic formalism.

Reasoning [22,73,38]. It should be possible to propagate policies based on the semantic relations between authorisation subjects, objects and access rights. Using ontologies, rules or a combination of both, it is possible to perform deductive reasoning and ab-
Fig. 3. Access Control Specification
ductive reasoning over access control policies [3, 8, 47, 50, 60, 93, 95]. In addition, a number of authors have proposed propagation strategies based on RDFS entailment [32, 41, 50, 69, 70] and hierarchies, partial orders or ontological relations between RDF resources [43, 65, 73, 86]. Kirrane et al. [52] propose a flexible authorisation framework which can be used to specify declarative authorisations, propagation policies, integrity constraints and conflict resolution rules. Unlike the other authors, who use reasoning to either infer or to propagate access control policies, Bao et al. [4] demonstrates how it is possible to reason over restricted data without releasing any restricted information. However, the interplay between the various reasoning strategies proposed, and the access control requirements arising from concrete use cases remains an open issue.

**Condition Expressiveness** [23, 8]. When it comes to access control, it should be feasible to specify conditions under which a request will be permitted or prohibited. The majority of the access control strategies that were examined support both authorisations and obligations. However, ODRL [86] and the general policy languages [3, 47, 93, 95] also catered for obligation and dispensation policies.

**Evidences** [22, 23, 8, 90]. According to The W3C Linked Data Platform Working Group [90] it should be possible to authenticate a subject (otherwise known as an agent) with an identifier or as the owner of a token. However, as the requester may be unknown to the system prior to submitting a request, access should be based on properties pertaining to the requester, commonly known as attributes, instead of traditional identities [38, 18, 32, 33, 47, 52, 56, 70, 73, 75, 93]. It should also be feasible to dynamically activate policies based on context [11, 38, 18, 32, 33, 47, 65, 75, 93, 95, 86]. Context can relate to the requester, the system or the environment. Attributes and context should be communicated by means of digital certificates, known as evidences. The de facto standard for submitting evidences is WebID [60, 18, 52, 75]. However, a number of researchers have also proposed using OpenID [19, 8, 60, 75]. According to Inkster et al. [40] although both WebID and OpenID serve the same purpose, WebID has a number of benefits over OpenID. For example WebID is simpler to use as a User Agent remembers the URI on behalf of the user, the establishing of identity is easy, WebID is truly decentralized and WebID is fully distributed. For additional details the reader is referred to the detailed comparison presented in Inkster et al. [40].

**Interoperability** [104]. Access control for open distributed environments, such as the web, needs to be able to support a wide variety of disparate policies, resources and users. One of the primary goals of standardisation is to maximize interoperability. As each of the access control strategies examined use open standards, such as RDF, RDFS, OWL and SPARQL, regardless of the specific use case they are suitable for access control over Linked Data. Using ontologies it is possible to specify access control vocabularies that can easily be adopted by others. In addition, OWL predicates such as `owl:sameAs` and `owl:disjointFrom` can be used to merge different access control vocabularies.

### 5.2. Enforcement

Access control enforcement requirements refer to constraints that are placed on the policy language or mechanisms that assist the requester to complete their request. An overview of the requirements is presented below and a snapshot of existing support for said requirements is presented in Table 6 and the corresponding timeline is represented in Figure 4.

**Information & Negotiation** [22, 23, 8, 90]. In order to protect user privacy, it should be possible for both the service provider and the requester to define policies and exchange credentials until an agreement has been reached. The W3C Linked Data Platform Working Group [90] indicate that service providers should describe access control policies for a given resource and user agents should be able to find the policies associated with a resource. The process is commonly known as negotiation. The P3P recommendation and the APPEL vocabulary have been designed to support automatic negotiation between clients and servers. The access control mechanisms proposed by Amini and Jalili [3], Bonatti and Olmedilla [8] and Toninelli et al. [93] all cater for access control negotiation.

**Explanations** [23, 8, 90]. According to De Coi et al. [23], Bonatti and Olmedilla [8], The W3C Linked Data Platform Working Group [90], service providers should provide explanations why access was not al-

---

19OpenID, http://openid.net/
Table 6

<table>
<thead>
<tr>
<th>Enforcement requirements</th>
<th>Information &amp; Negotiation</th>
<th>Explanation</th>
<th>Conflict Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abel et al. [1]</td>
<td>-</td>
<td>-</td>
<td>default</td>
</tr>
<tr>
<td>Amini and Jalili [3]</td>
<td></td>
<td>bidirectional policies</td>
<td>-</td>
</tr>
<tr>
<td>Ehsan et al. [26]</td>
<td></td>
<td>-</td>
<td>algorithm</td>
</tr>
<tr>
<td>Bao et al. [5]</td>
<td></td>
<td>-</td>
<td>algorithm</td>
</tr>
<tr>
<td>Bonatti et al. [10]</td>
<td></td>
<td>bidirectional policies</td>
<td>-</td>
</tr>
<tr>
<td>Bonatti and Olmedilla [73]</td>
<td></td>
<td>metarules &amp; queries</td>
<td>-</td>
</tr>
<tr>
<td>Chen and Stuckenschmid [44]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Villata et al. [110]</td>
<td></td>
<td>-</td>
<td>default</td>
</tr>
<tr>
<td>Costabellos et al. [14,15,20]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dietzold and Auer [25]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fournier et al. [31]</td>
<td></td>
<td>-</td>
<td>default</td>
</tr>
<tr>
<td>Franzoni et al. [42]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gabillon and Letouzey [33]</td>
<td></td>
<td>-</td>
<td>default</td>
</tr>
<tr>
<td>Jain and Parkas [34]</td>
<td></td>
<td>-</td>
<td>algorithm</td>
</tr>
<tr>
<td>Javanmardi et al. [41]</td>
<td></td>
<td>-</td>
<td>algorithm</td>
</tr>
<tr>
<td>Kagal et al. [45]</td>
<td></td>
<td>-</td>
<td>meta policies</td>
</tr>
<tr>
<td>Kagal et al. [46]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kagal and Bormes-lee [45]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Denker et al. [48]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kim et al. [49]</td>
<td></td>
<td>-</td>
<td>algorithm</td>
</tr>
<tr>
<td>Kirrane et al. [51,53,52]</td>
<td></td>
<td>-</td>
<td>flexible framework</td>
</tr>
<tr>
<td>Lopes et al. [56]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kolovski et al. [56]</td>
<td></td>
<td>-</td>
<td>priorities</td>
</tr>
<tr>
<td>Li and Cheung [77]</td>
<td></td>
<td>-</td>
<td>harmonisation</td>
</tr>
<tr>
<td>Mühliesen et al. [60]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Onilahghone et al. [63]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Papakonstantinos et al. [65]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Qin and Atluri [69]</td>
<td></td>
<td>-</td>
<td>default</td>
</tr>
<tr>
<td>Reddivari et al. [70]</td>
<td></td>
<td>-</td>
<td>meta policies</td>
</tr>
<tr>
<td>Ryutov et al. [72,73]</td>
<td></td>
<td>-</td>
<td>user interface</td>
</tr>
<tr>
<td>Sacco et al. [74,75]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sacco and Pattani [76]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sacco and Breslin [77]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Toninelli et al. [92,93]</td>
<td></td>
<td>bidirectional policies</td>
<td>descriptions</td>
</tr>
<tr>
<td>Strejskal and Pollet [78]</td>
<td></td>
<td>-</td>
<td>harmonisation</td>
</tr>
<tr>
<td>Usnik et al. [82,83,84,85]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bradshaw et al. [86]</td>
<td></td>
<td>-</td>
<td>harmonisation</td>
</tr>
<tr>
<td>Johnson et al. [87]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Suri et al. [88]</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

allowed. Rather than simply granting or denying access, the policy should also provide details of how the decision was reached. Such explanations would be of benefit to both the requester and the policy owner, making it easier for the requester to understand what is required of them and for the policy owner to troubleshoot potential problems. Bonatti and Olmedilla [8], Ryutov et al. [73] and Toninelli et al. [93] all provide policy explanations. However, both Ryutov et al. [73] and Bonatti and Olmedilla [8] provide a means to execute queries over policies in order to obtain additional information. It is worth noting that explanations should be used with care as such information could in fact weaken the security of the policy.

**Conflict Resolution [3].** Conflicts between both explicit and implicit policies should be resolved automatically. A number of different conflict resolution strategies have been proposed. In the event of a conflict some authors simply default to grant/deny [118,131,165,69,86], use priorities to determine dominance [50] or use metapolicies as a flexible means to resolve conflicts [47,70]. A number of authors propose conflict resolution algorithms based on several different measures [33,41,43,50,73]. Kirrane et al. [52] suggest a general syntax, which can be used to specify declarative conflict resolution policies. Whereas, others try to isolate individual data items that are in conflict and propose harmonisation strategies [57,93,55].
5.3. Administration

This section presents a number of access control requirements that are necessary to simplify the specification and maintenance of access control policies. An overview of the requirements is presented below and a summary of current support is presented in Table 7 and the corresponding timeline is represented in Figure 5. Although a number of researchers indicate that they provide some level of support for these requirements, generally speaking research efforts seem to focus more on the specification and enforcement mechanisms, and very little detail is supplied.
Table 7
Administration requirements

<table>
<thead>
<tr>
<th></th>
<th>Delegation</th>
<th>Consistency &amp; Safety</th>
<th>Usability</th>
<th>Understandability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abel et al. [1]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Amini and Jalili [3]</td>
<td>-</td>
<td>case study</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Elsan et al. [26]</td>
<td>-</td>
<td>algorithm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bao et al. [5]</td>
<td>-</td>
<td>algorithm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bonatti et al. [10]</td>
<td>-</td>
<td>language</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bonatti and Olmedilla [72]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Protunex explanation</td>
</tr>
<tr>
<td>Chen and Stuckenschmidt [15]</td>
<td>-</td>
<td>meta policies</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Villata et al. [100]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Costabello et al. [73,74,93,90]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dietzold and Auer [25]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Floris et al. [31]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Franzoni et al. [32]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gabillon and Letouzey [33]</td>
<td>-</td>
<td>construct &amp; describe queries</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Jain and Farkas [41]</td>
<td>-</td>
<td>algorithm</td>
<td>-</td>
<td>RACL admin module</td>
</tr>
<tr>
<td>Jawannardi et al. [33]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kagal et al. [47]</td>
<td>-</td>
<td>speech acts</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kagal et al. [48,49]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kagal and Berners-lee [46]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Denker et al. [24]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kim et al. [50]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kirrane et al. [51,53]</td>
<td>-</td>
<td>DAC</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lopes et al. [53]</td>
<td>-</td>
<td>flexible framework</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kolovski et al. [56]</td>
<td>-</td>
<td>analysis services</td>
<td>-</td>
<td>analysis services</td>
</tr>
<tr>
<td>Li and Cheung [37]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mühleisen et al. [60]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ouelhadjmoune et al. [45]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Papakonstantinou et al. [52]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Qin and Athari [20]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reddittari et al. [70]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ryutov et al. [72,75]</td>
<td>-</td>
<td>meta policies</td>
<td>RAW policy editor</td>
<td>RAW permission check</td>
</tr>
<tr>
<td>Sacco et al. [74,75]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sacco and Passant [76]</td>
<td>-</td>
<td>privacy manager</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sacco and Breslin [74]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Steyskal and Politzer [57,58,23]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Toninelli et al. [24,24]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bradshaw et al. [11,12]</td>
<td>-</td>
<td>K AoS policy admin tool</td>
<td>policy disclosure</td>
<td></td>
</tr>
<tr>
<td>Usnik et al. [55,56,75,80,82]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Johnson et al. [44]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Suri et al. [39]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Delegation [8]. It should be feasible to temporarily transfer access rights to other users. In relational databases, users are granted sole ownership of the tables and views that they create. They can subsequently grant access rights to other database users. Amini and Jalili [3], Bonatti and Olmedilla [8], Gabillon and Letouzey [33] and Kagal et al. [47] indicate that they support the delegation of access rights. Kirrane et al. [52] demonstrate how the discretionary access control model can be used to guide access control specification and administration. However, the suitability of existing revocations strategies, for the RDF graph model, warrants further research.

Consistency & Safety [73]. In order to ensure the access control system is complete and accurate, insertion and deletion of policies should be controlled. It should not be possible to elevate your own privileges or to assign permissions that would make data inaccessible to everyone. Although a number of researchers indicate that their frameworks support consistency and safety constraints, very little information is provided. Kolovski et al. [56], Amini and Jalili [3], Bao et al. [4] and Jain and Farkas [41] ensure consistency and safety as part of their administration algorithms. Chen and Stuckenschmidt [14] and Ryutov et al. [75] suggest that metapolicies can be used to ensure consis-
tency and safety. Kirrane et al. [52] propose a general syntax for integrity constraints, which can be used to specify declarative constraints. Given the diversity of access control models, policies and reasoning strategies that have been proposed for RDF, additional research is required in order to determine potential issues with access control policies and propose suitable handling mechanisms.

Usability [22][3]. The specification and the maintenance of access control policies should be as simple as possible. Administration facilities that support ease of both specification and maintenance of policies have been provided by a number of researchers [41][73][6][95][20]. Jain and Farkas [41] provide a screenshot of their RACL admin tool, however no explicit details are provided with respect to the functionality. According to Costabello et al. [20], Uszok et al. [95], Ryutov et al. [73], Sacco and Passant [76] their policy managers proposed by enable administrators to create, edit and deleting access control policies. However, it is not clear how usable these interfaces are when it comes to the administration of complex policies over large datasets. Given the complexity associated with reasoning over graph data, advanced data analytics and visualisation techniques are needed to highlight the effects of advanced policies, constraints and deduction rules.

Understandability [73][3]. It should be easy to understand the interplay between policies. Only a handful of researchers associate policy explanations with policies. Similarly, only a select few provide systems that enable administrators to verify the interplay between policies [8][56][3][95]. Given the dynamic nature of context based access control and the various deduction and propagation strategies, further research on automating the explanations and presenting the results in a manner which is digestible by humans is necessary.

5.4. Implementation

Implementation requirements generally refer to non-functional requirements. As with any software system, non-functional requirements hold the key to the adoption of a tool or technology. Although a number of authors indicate that the solutions they propose are flexible or extensible, seldom do researchers evaluate these claims. Table 8 and the corresponding timeline is represented in Figure 6 provide an overview of the technologies adopted and indicates the evaluations performed. Where applicable a link to the ontology, framework, dataset or demo is provided under the heading 'Available Resources'.

Effectiveness [73]. In order to work in practice, access control enforcement and administration needs to be efficient. A number of authors have presented performance evaluations of their access control enforcement [4][8][3][56][2][6][7][8], query rewriting [1], annotation [3], explanation [8] or reasoning [4][56][6][5] algorithms. In general, the authors reported a linear increase in performance over increasing policies. Both [4][56] highlight the fact that reasoning can be expensive and consequently only allowing subsumption relations with ontology concepts. The most comprehensive evaluation is performed by [8]. The authors automatically generate policies for both real and artificial scenarios in order to evaluate the performance of their explanation and negotiation modules. In addition, they discuss the expressiveness of the Protune policy language. However, there is still no clear access control benchmark, that can be used to compare different approaches in terms of policy enforcement and administration. One suggestion would be to build a set of access control scenarios and extend the BSBM dataset generator to cater for solution benchmarking (i.e. auto generate access control policies and constraints over increasing datasets).

Distributed [3]. In order to ensure scalability, it should be possible to cater for the distributed specification and enforcement of access control policies. A number of researchers have examined how their proposed solution can be applied to use cases requiring distributed access control mechanisms. Amini and Jalili [3] describe a case study on distributed semantic digital library. Bonatti and Olmedilla [8] and Mühleisen et al. [60] developed demos in order to demonstrate how their policy languages can be used in a distributed setting. Whereas, the access control languages and en-
Table 8

<table>
<thead>
<tr>
<th>Implementation requirements</th>
<th>Effectiveness</th>
<th>Distributed Use case</th>
<th>Availability Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abel et al. [1]</td>
<td>query rewriting performance</td>
<td>-</td>
<td>SeRQL, Protune</td>
</tr>
<tr>
<td>Amini and Jalili</td>
<td>enforcement performance</td>
<td>case study</td>
<td>Prolog, GWT, Jena, JIP, Protege</td>
</tr>
<tr>
<td>Bao et al. [2]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bonatti et al. [10]</td>
<td>explanation performance</td>
<td>demo</td>
<td>Java, TuProlog</td>
</tr>
<tr>
<td>Chen and Stuckenschmidt [4]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Costabello et al. [18]</td>
<td>-</td>
<td>mobile use case &amp; demo</td>
<td>SAAC Ontology</td>
</tr>
<tr>
<td>Dietzold and Auer [22]</td>
<td>-</td>
<td>-</td>
<td>RDF, SPARQL</td>
</tr>
<tr>
<td>Flower et al. [31]</td>
<td>-</td>
<td>-</td>
<td>Java, Sesame, Progress</td>
</tr>
<tr>
<td>Franzoni et al. [33]</td>
<td>-</td>
<td>-</td>
<td>Java, SeRQL, Sesame</td>
</tr>
<tr>
<td>Gabrielsson and Leksand [35]</td>
<td>-</td>
<td>Java, Tomcat, Corese-KGRAM, Corese</td>
<td></td>
</tr>
<tr>
<td>Jain and Parkar [41]</td>
<td>-</td>
<td>-</td>
<td>Java, Jena, Jess</td>
</tr>
<tr>
<td>Javaamidari et al. [43]</td>
<td>-</td>
<td>-</td>
<td>PELLET, SWRL</td>
</tr>
<tr>
<td>Kagal et al. [47]</td>
<td>-</td>
<td>-</td>
<td>Java, Prolog</td>
</tr>
<tr>
<td>Kagal et al. [48]</td>
<td>-</td>
<td>-</td>
<td>Java, Prolog</td>
</tr>
<tr>
<td>Kagal and Breslin [49]</td>
<td>-</td>
<td>-</td>
<td>Java, Prolog</td>
</tr>
<tr>
<td>Denker et al. [50]</td>
<td>-</td>
<td>-</td>
<td>Java, Prolog</td>
</tr>
<tr>
<td>Kim et al. [51]</td>
<td>-</td>
<td>-</td>
<td>Java, Prolog</td>
</tr>
<tr>
<td>Kirrane et al. [52]</td>
<td>-</td>
<td>Linked Data use case</td>
<td>Linked Data use case</td>
</tr>
<tr>
<td>Lopes et al. [53]</td>
<td>-</td>
<td>-</td>
<td>Java, Jena, GFAF datasets</td>
</tr>
<tr>
<td>Kolovski et al. [54]</td>
<td>-</td>
<td>-</td>
<td>Pellet</td>
</tr>
<tr>
<td>Lü and Cheung [55]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mühleisen et al. [56]</td>
<td>-</td>
<td>-</td>
<td>Joseki, Jena, Pellet</td>
</tr>
<tr>
<td>Oulmachouane et al. [57]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Papanikolaou et al. [58]</td>
<td>-</td>
<td>-</td>
<td>PostgreSQL, MonetDB</td>
</tr>
<tr>
<td>Qin and Adami [59]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reddivari et al. [60]</td>
<td>-</td>
<td>-</td>
<td>Java, Jena, RDQL</td>
</tr>
<tr>
<td>Ryu et al. [61]</td>
<td>-</td>
<td>-</td>
<td>Java, Jena</td>
</tr>
<tr>
<td>Sacco et al. [62]</td>
<td>-</td>
<td>mobile use case</td>
<td>Java, PPO Ontology, PPM Scratchpad</td>
</tr>
<tr>
<td>Sacco and Passant [63]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sacco and Breslin [64]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Steyskal and Polleres [65]</td>
<td>-</td>
<td>Linked Data use case</td>
<td>-</td>
</tr>
<tr>
<td>Tominelli et al. [66]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Uszok et al. [67]</td>
<td>-</td>
<td>-</td>
<td>Proteus framework</td>
</tr>
<tr>
<td>Bradshaw et al. [68]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Johnson et al. [69]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Suri et al. [70]</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Forcement frameworks proposed by Costabello et al. [18], Kagal et al. [47], Sacco and Passant [63] and Uszok et al. [67] are motivated by distributed uses cases. Although Steyskal and Polleres [65] present Linked Data use cases, the adaptation of current distributed query processing techniques to cater for access control over Linked Data has not been explored to date.

**Flexibility & Extensibility**. The system should be capable of handling frequent changes to policies, user, access rights and resources. In addition, in order to provide support for different scenarios and future enhancements, the enforcement frameworks should be flexible and extensible. As each of the access control strategies examined use one or more open standards (see Table 5), they are by design flexible and extensible. An overview of the technologies used in each of the access control proposals examined is presented in Table 8.
6. Conclusions and Future Work

This paper provided an overview of relevant access control models (MAC, DAC, RBAC, VBAC, ABAC, CBAC) and standardisation efforts (XACML, WebID, WAC, P3P, APPEL, ODRL), and described how they have been either enhanced by/applied to RDF. A number of well known policy languages, that adopt ontology based, rule based and combined ontology and rule based access control enforcement mechanisms were examined in detail. Several different strategies that have been used to specify access control over RDF (triple patterns, views, named graphs and ontologies) and various reasoning, filtering and query rewriting strategies were presented. Finally, a set of requirements for Linked Data, based on several papers that examine access control for RDF from a number of perspectives, were derived. These requirements were subsequently used to classify the various access control specification, enforcement and administration strategies that have been proposed for RDF data.

Based on this analysis a number of gaps with respect to access control for Linked Data, which still need to be addressed were identified:

Usability & Understandability. Access control administration in general, and over large datasets in particular, can become extremely difficult to manage. Access control policies may be composed of authorisations specified at multiple levels of granularity. In addition, permissions may be inferred or propagated using different inferencing mechanisms, making the task of administration is even more cumbersome. An interesting avenue for future work, would be to investigate if graph based data clustering and visualisation techniques, such as those proposed by [61], can be used to assist systems administrators to examine the interplay between authorisations and rules, and also determine the impact of new authorisations.

Explanations & Negotiation. The benefits associated with explanations are two fold: (i) they allow the requester to understand what is required of them and (ii) they enable the policy owner to troubleshoot potential issues with existing poli-
Effectiveness. In order to work in practice, both access control enforcement and administration need to be effective from both a performances and a correctness perspective. Although a number of authors have conducted access control performance evaluations using the BSBM dataset, when it comes to access control for RDF data there is currently no general access control benchmark. In addition, there is a pressing need for general mechanisms that can be used to verify the correctness of proposed access control strategies.

References


[30] Timothy Finin, Anupam Joshi, Lalana Kagal, Jianwei Niu, Ravi Sandhu, William Winsborough, and Bhavani Thuraisingham. ROWL.BAC: representing role based access control in OWL. In Indrakshi Ray and Ninghui Li, ed-


Sara Javanmardi, Morteza Amini, Rasool Jalili, and Yaser Ganjisaffar. SBAC: “a semantic-based access control model”. In Viiveke Fäk, editor, 11th Nordic Workshop on Secure IT-systems (NomSec@Årø), Linköping, Sweden, volume 22, 2006.


