



**HAL**  
open science

# IdSM-O: an IoT Data Sharing Management Ontology for Data Governance

Nouha Laamech, Manuel Munier, Congduc Pham

► **To cite this version:**

Nouha Laamech, Manuel Munier, Congduc Pham. IdSM-O: an IoT Data Sharing Management Ontology for Data Governance. MEDES '22: International Conference on Management of Digital EcoSystems, Oct 2022, Venice, Italy. pp.88-95, 10.1145/3508397.3564825 . hal-04206735

**HAL Id: hal-04206735**

**<https://univ-pau.hal.science/hal-04206735v1>**

Submitted on 14 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# IdSM-O : an IoT data Sharing Management Ontology for Data Governance

Nouha Laamech  
laamech.nouha@univ-pau.fr  
Universite de Pau et des Pays de  
l'Adour, E2S UPPA, LIUPPA  
Mont-de-Marsan, France

Manuel Munier  
manuel.munier@univ-pau.fr  
Universite de Pau et des Pays de  
l'Adour, E2S UPPA, LIUPPA  
Mont-de-Marsan, France

Congduc Pham  
congduc.pham@univ-pau.fr  
Universite de Pau et des Pays de  
l'Adour, E2S UPPA, LIUPPA  
Pau, France

## Abstract

The main purpose of IoT is to deliver reliable, high quality services and innovative solutions by transforming the captured data into meaningful information, and thus improving user's daily life. In this regard, it is in the interest of the community to encourage entities within IoT environments to share their data, and therefore serve public interest and contribute to the innovation and technological progress. Meanwhile, the distributed nature of IoT networks and the diversity of its actors lead to the recognition of security and data sharing management as one of the major challenges of the IoT domain. For instance, due to insufficient governance of the shared data within IoT environments, data provider retains little to no control over his assets once he has agreed to share them. Furthermore, data consumers are not able to trace the source of the available resource nor its history processing to assess its quality. All this creates a digital environment that is certainly functional but lacks mutual trust between its actors, which can prevent the domain's full potential to be exploited, and therefore disrupt the implemented services. In our work, we propose an approach to improve data sharing management using three main elements: semantic modeling, usage control policies, and data provenance.

**CCS Concepts:** • Security and privacy → Trust frameworks; • Information systems → Data management systems.

**Keywords:** data sharing management, usage control, semantic modeling, data governance

## ACM Reference Format:

Nouha Laamech, Manuel Munier, and Congduc Pham. 2022. IdSM-O : an IoT data Sharing Management Ontology for Data Governance. In *Proceedings of the 14th International Conference on Management of Digital EcoSystems (MEDES '22)*. ACM, Venice, Italy, 8 pages. <https://doi.org/10.1145/3508397.3564825>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MEDES '22, October 19–21, 2022, Venice, Italy

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9219-8/22/10...\$15.00

<https://doi.org/10.1145/3508397.3564825>

## 1 Introduction

The IoT constitutes both a technological and strategic transition, which promises to improve everyday life and enhance its quality. Numerous mainstream devices, which have been inactive until now, will be able to collect, analyze and share available observation on behalf of their owners. At the same time, data consumers can reuse this shared data and provide in return several new high quality services in various domains such as smart agriculture, health care, etc.

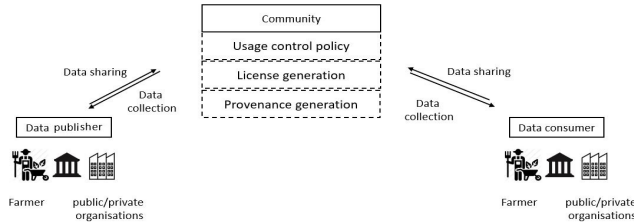
In this context, we can see that it is in the best interest of the community to encourage and incite connected environments actors to share their data as much as possible. Therefore, other entities could access it, and thus contribute to technological progress, innovation and improve the daily life of users.

In some cases, users agree to share their resources within Open Data. Open data [1] consists in making available to the community, free of charge, the largest possible volume of data. However, Open Data as it is conceived today is far from being a resilient and full range solution. The inability to trace the source of the asset as well as its processing history makes it impossible for the data consumer to assess its quality and decide whether it's adequate for their business needs or not. Furthermore, from the moment the data is deployed, there is no way to guarantee that the data owner will stay updated on how his data is being used. Those challenges make it difficult for the available assets to be exploited and creates an environment that lacks trust between its actors.

## 2 Foundational material

### 2.1 Motivating example

Farmers capture several types of data related to their particular land and business operations, such as irrigation and soil moisture levels, vineyard structures, etc. By sharing this data with the community, farmers would be able to compare their production with that of their neighbors, while still pursuing their own specific business objectives. Meanwhile, other autonomous organizations will also have access to these information in order to provide a range of innovative services to help the community, such as water management, resources monitoring, or yet review waste and pollution control procedures and therefore accelerate the agro-ecological transition.



**Figure 1.** Use case of information security enforcement

We illustrate with this case study the issues related to data sharing management and how to use usage control policy to specify information security requirements. Within smart agriculture environment, inciting farmers and different entities to share their collected data with the rest of the community allows to improve advanced services across a large scale of application domains (Figure 1). However, we can clearly see how IoT systems are open communities without much standardized monitoring. Data providers don't have the ability to stay informed where their data is being distributed and in which processes it's being involved. Meanwhile, data consumers don't have the right tools to trace the source of the available resources, its processing history, and whether they have the right to use it for their goals. All this disrupts the proper functioning of IoT environment and creates lack of mutual trust.

## 2.2 Data sharing considerations

Based on current data management concerns, we can identify the following key concepts of our information security system:

- The identification of data producer preferences: data producers set the requirements that need to be respected to be allowed to use their assets. The requirements correspond to a set of OrBAC rules, that help data owners take advantage of their status without the need to have enough experience and expertise in the security domain.
- The license generation: a semantic license is generated based on the previously settled preferences and is assigned to the IoT data, in such a manner that regardless of where the data is located, the license will stay with it.
- Data governance: our work joins the "Data Governance Act" (DGA) [3] of the European Commission, which aims to promote data sharing by creating a digital environment of trust between different actors. This vision can be applied to various fields: health, energy, smart

cities, etc.

- Common good: our solution promotes an IoT environment where different resources can be shared within communities to benefit the general interest [4].

## 3 Contextual usage control management

Incorporating information security requirements into a security policy, by translating it into a usage control model, will enable a convenient upgrade to existing information systems that have already established usage control policies.

### 3.1 Motivation for using OrBAC Model

Organization-Based Access Control model (OrBAC) [9] is a contextual usage control model that provides mechanisms to express security policies, and enables distinction between an abstract policy specifying organizational requirements and its implementation in a given information system.

The abstraction of traditional access control entities (subject, action, object) into meta entities (role, activity, view) allows the development of a security policy at two levels, a concrete level and an abstract level [5]. The introduction of an abstract organizational level also allows for a structure as shown in Figure 2 [13]. We obtain a two-level security policy. As a result, the OrBAC model makes it possible to establish an abstract security policy (role, activity, view) independent of the implementation choices (subject, action, object).

Safety rules that can be used at the abstract level have the following form:

*permission(org,role,activity,view,context)* : indicates that in the organization *org*, *role* is allowed to perform an activity on the view if the context is valid.

*prohibition(org,role,activity,view,context)* : has the same parameters as the permission predicate but declares a prohibition for a role to do an activity on a view in an organization when a context is valid.

*obligation(org, role, activity, view, context, violationCtx)* : unlike the predicates of permission and prohibition, an obligation has an extra context, the context of violation, which specifies the condition in which the obligation is violated.

By using this model, each organization can define security rules to specify whether certain roles are allowed, prohibited or required to perform certain activities on certain views. The activation of those rules depends on contextual conditions rather than using static ones. Since a security policy may include conflicting security rules (e.g. conflict between a permission and a prohibition), it is possible to define conflict resolution strategies based on the assignment of priority to rules.

The OrBAC model has some valuable features:

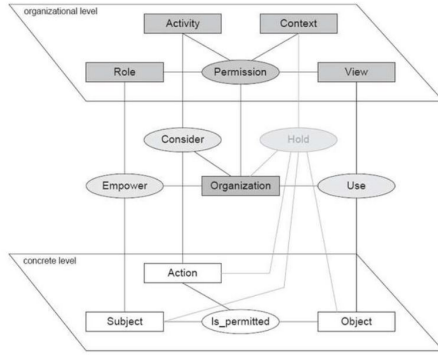


Figure 2. OrBAC model (taken from [13])

- The concept of context is introduced in the model and a possibility to express explicit contextual authorizations. Contexts could be activated or not depending on defined instances, which could be helpful to eventually take into consideration data provenance, and therefore establish a dynamic security policy by triggering certain OrBAC rules depending on the context of an activity.
- The establishment of a security policy within the organization. This reflects the general application of privacy practices to an organization over which data owners define their preferences. They also outline the obligations that requesters must fulfill after accessing sensitive information.  
For those reasons, we chose the OrBAC model to establish a usage control policy for smart agriculture.

### 3.2 Modeling a usage control model within OrBAC

In order to implement usage control policies for data sharing management in smart agriculture, we intend to illustrate how it can be specified through the OrBAC model.

We define an organization as the "Data\_Producer\_Facility", where the data provider is the only authority over the different actors that play a role within the structure. By doing this, each data producer node is responsible for making his own decision.

The subjects within a Data Producer Facility correspond to users of this network. The assignment of access rights to these topics is done through role structuring. Subjects are assigned permission rights to observations shared by a data producer, by playing roles. A role corresponds to a set of usage control rules and is only meaningful in the organization where it has been assigned. Let 'farmer\_alice' be a data producer in the organization 'Data\_Producer\_Facility', and 'water\_management\_organization' a data consumer. The data producer is considered as a subject within the organization. We define:

- The role 'data\_provider'  
*Empower("Data\_Producer\_Facility","farmer\_Alice", "data\_provider")*
- The role 'data\_consumer'  
*Empower("Data\_Producer\_Facility", "water\_management\_organization","data\_consumer")*
- Two sub-roles: 'professional' and 'individual':  
*Sub-role("Data\_Producer\_Facility", "professional", "data\_consumer")*  
*Sub-role("Data\_Producer\_Facility", "individual", "data\_consumer")*
- The activity 'access'
- The view 'wind\_speed\_observation'

We define the permissions that enforce privacy policy within the Data\_Producer\_Facility organization for our use case. The first permission allows 'professional' to access wind speed observation for marketing purposes :

```
Permission("Data_Producer_Facility","professional","access",
"wind_speed_observation","marketing_purposes");
```

The second permission allows 'individual' to access 'wind\_speed\_observation' if the data consumer delete the data provider right after he finish his treatment with it:

```
Permission("Data_Producer_Facility","individual","access",
"wind_speed_observation","delete_after_finished_treatment")
```

## 4 IoT Data Sharing Management Ontology overview

The IoT is an evolving, interactive environment in which smart devices interact with each other to execute high-level tasks. We chose to describe IoT data using ontology in order to abstract implementation issues and allow us to establish basis concepts for describing security aspects. Furthermore, using ontological data model allows us to address interoperability of various concepts used by multiple IoT actors in heterogeneous domains to define shared data. The main goal of our ontology is to build a data-sharing management system that preserve data provider's important role throughout the process of collecting, transmitting, storing, and processing data collected by smart devices. We use standard ontology languages to define a common privacy vocabulary, combined with standard reasoning technologies based on description logic to meet data governance requirements. Moreover, in order to achieve interoperability to the fullest extent possible, our ontology imports the Semantic Sensor Network (SSN) ontology [8], PROV-O [11] and the OrBAC ontology [12] to

specify certain classes and extend it with certain information security properties. SSN and prov-O are published and recommended by the World Wide Web Consortium (W3C). Although they present knowledge in the domain of sensor networks, data provenance, and usage control, those ontologies lack the relevant definitions to enhance a good data sharing management.

The first task is to conduct an alignment of these three ontologies and their concepts. To be compliant with today ontology engineering, we adopt the modularization method, which consists in segmenting an ontology into smaller parts. Figure 3 depicts some features of the different IdSM-O modules. In order to cover information security requirements, IdSM-O contains three main components, namely IoT description module, IoT data sharing management module, and IoT data provenance. Each module includes a set of concepts. It aims at providing the users with the knowledge they require, narrowing the scope as much as possible to what is strictly necessary in a given use case.

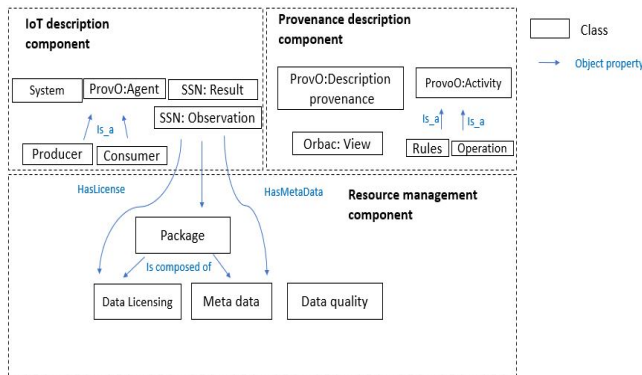


Figure 3. IdSM Ontology overview

#### 4.1 IoT description component

The IoT description component describes the IoT environment. Therefore, we will focus on building our IoT description model based on the the following scenario :

"When measuring the temperature of a room, the temperature is the observation, 26 degree can be the result of the observation and the room is the FeatureOfInterest.

Although basic, this scenario will help us identify the main entities within a typical IoT environment, and therefore be able to build a simple yet consistent model. In this context, our IoT description component consists of three main sub-modules, namely:

- Agents : which includes the different parties and collaborators involved to achieve a common goal in the IoT domain.
- IoT resource: which focus on modeling the observation, the feature of interest, and the output result.

- System: which describes system networks, including community, sensor, and organization.

We reuse some conceptual constructs from the SSN ontology and extend it with new classes to facilitate access, use, and auditing verification of information generated by IoT resources. Below, we detail the second module, which aims to describe the provenance of an observation.

#### 4.2 Provenance description component

The provenance description module helps the data consumer to have a better visibility on where the requested observation comes from, what generated it, and what invalidate it. Hence, the data consumer is able to assess the observation quality while taking in consideration his own goals and own subjective definition of "data quality".

In the same approach as SSN, we reuse some PROV-O classes, subclasses, and properties to build our provenance model. This module includes the following classes:

- Activity class: this class defines the actions occurring within the system.
- OrBAC:View class: an element that has some static aspects. In this system, prov:Entity would be aligned with OrBAC:View ( $ssn : Observation \in prov : Entity$ ).
- Rules: contains three classes : is\_Obligated, is\_Permitted, and is\_Prohibited. Describes the permitted, prohibited, and eventually the obligated actions on an observation.
- Description provenance class: an URI that permit to navigate thought a knowledge graph. Contains provenance description of an observation, and is itself a View, so allowing provenance of provenance to be expressed.

#### 4.3 Resource management component

The IoT Resource Management module helps the owner to better control his IoT resources and not lose his ownership once his asset is shared. Hence, this module includes two sub-modules, namely:

1. Owner's Information security preferences sub-module: this sub-module helps the owner to define the information security to preserve the ownership of his provided data and to express his preferences on how his smart devices must behave. It includes Rules, Activity, and View classes.
2. License and package generating sub-module: data owner defines OrBAC rules that describe how he want his data to be handled when it's being shared. A semantic license is then generated based on those preferences and is attributed to the IoT data. Therefore, when a data consumer request an observation, he describes

with basic OrBAC rules how he intend to use it, and the system decide if data owner's terms of use match data requester use description using SWRL reasoning, to finally conclude whether he has the right to access the observation or not, or the terms that he needs to agree on before using it.

- a. "Data\_licensing" class: contains SWRL rules that translate data owners preferences from OrBAC rules using semantic reasoning.
  - b. "Metadata\_description": a file that contains metadata related to an observation. It contains provenance information class, which is itself a View, allowing therefore provenance of provenance to be expressed.
  - c. "Package" class: a feature consisting of a raw observation, its generated license, and its metadata description. Therefore, when an observation is broadcast in an IoT environment, the entire packet is transmitted, rather than just the data.
3. Data quality sub-module: by improving metadata and provenance models, data consumers are allowed to make better sense of the available data, and therefore have a better visibility to assess their data quality. We define 5 subclasses : Accuracy, Frequency, Granularity, Precision, and Timeliness[10]. Note that these are some measures examples, and that we can still modify them.

## 5 Reasoning process

Reasoning is an important inherent function of ontology, and reasoning rules can be added as a part of the defined ontologies to infer the information implied into them. In this work, the ontology previously developed and the different descriptions model are respectively regarded as the basis for reasoning. SWRL [7] is used as the chosen tool to define the reasoning rules needed to implement mutual understanding and interactions between observations and the involved actors.

### 5.1 Smart agriculture case study

We illustrate our approach in a smart agriculture context, but our ontology is agnostic and can be applied in other IoT contexts. Thus, we describe below a motivating smart agriculture scenario:

Bob is a farmer who has web-enabled smart devices around his lands that use embedded systems, such as sensors, to collect, send and act on data. This allows Bob to enhance productivity, monitor his business processes, and improve his revenue. The sensor collects these data and sends them to the data-provider-gateway through a secure channel. In addition, the data-provider-gateway enables Bob to adjust his device settings, including permission and usage control,

Atom	Description
Observation(?ob)	Ob is an instance of Observation
hasView(?ac, ?ob)	Activity ac is applied on the view ob, which is an observation.
hasRequestedObservation(?dc, ?ob)	A data consumer dc had requested to use the observation "ob"
hasSubject(?ac, ?ag)	The activity ac has subject an agent ag
hasOperation(?dc, Write_operation)	The data consumer "dc" has "write operation" rights
hasContext(?ac, ?cntx)	An activity ac has a cntx context.

**Table 1.** Atoms description

and to establish the conditions that need to be respected if a service provider decide to use his available data. From the water management center, a public organization can remotely access and monitor water usage by receiving Bob's devices observation. When receiving those resources, the water management access a package that contains the observation, its license, and its metadata information. The license contains the conditions data consumers need to respect while using the observation, and the metadata contains it provenance information and its data quality. In case the organization agrees to them, it can start using it to provide mechanisms for water management to avoid shortages and avoid water wastage.

### 5.2 SWRL Atoms and licenses generation

While OWL is used to implement the IdSM-O classes, OrBAC rules such as those defined in section 3 cannot be expressed in OWL. For this purpose, we use the Semantic Web Rule Language (SWRL) to construct a set of inference rules, which are built on different concepts and properties of IdSM-O. The inference rule represents a set of conjunctions of atoms, called an antecedent that implies a result, called a consequent. Thus, based on SWRL rules and IdSM-O classes, multiple semantic licenses can be inferred for different possible data sharing cases in the real world. We propose below the set of inference rules.

#### A) SWRL Atoms

According to SWRL language standards and the IoT data sharing management ontology created at section 4, the partial atoms of SWRL rules based on IdSM ontology is listed in Table1.

## B) SWRL Rules

Semantic licenses are generated during the reasoning process to derive requirements that data consumers need to agree on to be able to use an observation handled by a data owner.

Instantiating a Data Licensing class is the result of a successful translation from the data owner's OrBAC rules to SWRL rules using the defined processing attributes. Let's take the following OrBAC rules :

$R1 \equiv \text{Permission}(\text{"Data\_Producer\_Facility"}, \text{"farmer"}, \text{"access"}, \text{"soil\_humidity"}, \text{"delete\_after\_treatment"})$

$R2 \equiv \text{Permission}(\text{"Data\_Producer\_Facility"}, \text{"researcher"}, \text{"access"}, \text{"temperature"}, \text{"research\_purposes"} \& \text{"save\_in\_log"})$

The first rule states that a subject that has the role "farmer" can access the observation "soil\_humidity" if he deletes it immediately after treatment. R2 describes that a researcher can access the observation "temperature" for research purposes, on condition to save usage details in a log. Therefore, an obligation is triggered when the observation is accessed. The obligation specifies that a log data usage must be kept.

$R3 \equiv \text{Obligation}(\text{"Data\_Producer\_Facility"}, \text{"researcher"}, \text{"write"}, \text{"log"})$

Using the rules described in OrBAC in section 3 and the SWRL description used in section 5, we can translate the information security preferences stated by data owners into SWRL rules as follows. The terms in bold represent instances to be customized according to the preferences established by data owner, others will only be instantiated in data consumer's gateway, which is the reason we talk about licenses and not contracts. The license outcome act differently depending on who uses them. Figure 4 represents the translation of R1 from OrBAC rules to SWRL rules, and Figure 5 represents the translation of R2 and R3.

## 6 The system architecture

Figure 4 shows our proposed architecture that enables data sharing management using the previously defined ontology. The architecture involves two main entities, namely data consumer and data provider. The data provider uses his own portal to create and send his access request to the data provider portal. The data owner uses his portal via his gateway to define his preferences about each device observation output. During the reasoning process, the semantic rule engine evaluates the license generation by following

Consumer(?c)	$\wedge$ Observation(?obs)	$\wedge$
Professional(?r)	$\wedge$ Prov : Activity(?a)	$\wedge$
Context(?cntx1)	$\wedge$ IoT_device(?device)	$\wedge$
isMadeBy(?obs, ?device)	$\wedge$ OperationType(?op)	$\wedge$
hasRole(?c, ?r)	$\wedge$ hasView(?a, ?obs)	$\wedge$
hasSubject(?a, ?c)	$\wedge$ hasOperationType(?a, ?op)	$\wedge$
hasContext(?a, ?cntx1)		$\wedge$
sameAs(?obs, <b>soil_humidity</b> )		$\wedge$
sameAs(?cntx1, <b>delete_after_treatment</b> )		$\wedge$
sameAs(?r, <b>farmer</b> )	$\wedge$ sameAs(?op, <b>access</b> )	
$\rightarrow$ is_Permitted(?a)		

Figure 4. SWRL rule of R1

Prov:Activity(?a)	$\wedge$ Observation(?obs)	$\wedge$
IoT_device(?device)	$\wedge$ IoT_device(?device)	$\wedge$
Consumer(?c)	$\wedge$ Professional(?r)	$\wedge$ hasRole(?c, ?r)
OperationType(?op)	$\wedge$ isMadeBy(?obs, ?device)	$\wedge$
hasSubject(?a, ?c)	$\wedge$ hasOperationType(?a, ?op)	$\wedge$
hasView(?a, ?obs)	$\wedge$ sameAs(?obs, <b>temperature</b> )	$\wedge$
sameAs(?r, <b>researcher</b> )	$\wedge$ sameAs(?op, <b>access</b> )	$\wedge$
swrlx : makeOWLThing(?a, ?b)	$\wedge$ swrlx : makeOWLThing(?obs, ?log)	$\wedge$
$\rightarrow$ is_Permitted(?a)	$\wedge$ hasSubject(?b, ?c)	$\wedge$
hasView(?b, ?log)	$\wedge$ Log(?log)	$\wedge$
hasOperationType(?b, ?op)	$\wedge$ OperationType(?op)	$\wedge$
sameAs(?op, <b>write</b> )	$\wedge$ is_Obligated(?b)	

Figure 5. SWRL rule of R2 and R3

the usage control described in section 3 using our predefined set of SWRL rules. In case of a match between the data usage request and the data provider rules, a semantic license is created and sent to the data requester portal with the observation, which will allow actions on data only if it is mentioned in the license. Thus, our architecture will allow the deployment of a fully distributed infrastructure, without having to rely on a third party, and thereby allowing data circulation from one portal to another.

The semantic Rule Manager includes five core components, which are (i) OrBAC rules, the associated security rules to the data provider's preferences (ii) inference rules, which are the rules translated from OrBAC that enable matching data provider preferences with a license, (iii) rule engine, which is responsible for reasoning about the received Information Security and, then tacking a decision to create a License, (iv) query engine, which enables the rule engine to interrogate the IdSM ontology, and (v) IdSM ontology, which includes the various concepts and properties introduced in the previous section.

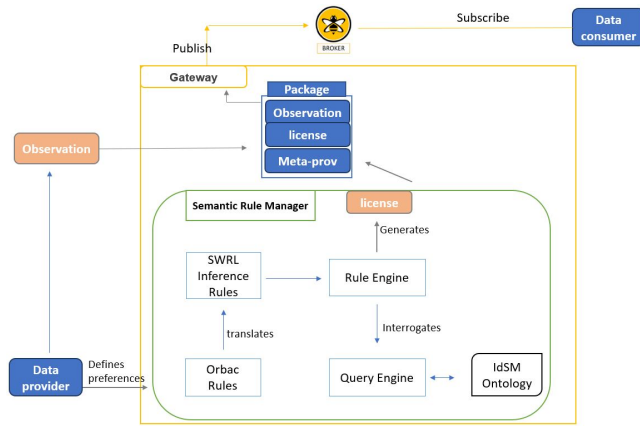


Figure 6. IdSM Ontology overview

## 7 Related work

Existing research approaches use the expressive strength of semantics and ontologies to maintain well-defined control within IoT environment.

The Semantic Sensor Network (SSN) ontology [8] was published in 2009, updated and recommended by the World Wide Web Consortium (W3C) in 2017. SSN describes sensors, actuators, samplers, and observations which are data collected by these sensors. It also provides a lightweight but self-contained core module, called SOSA (Sensor, Observation, Sample, and Actuator) [6]. Although the SSN ontology was able to model the domain of sensor networks and has become since then one of the major IoT models, it was short on the relevant definitions related to data management and security questions. Furthermore, it did not provide the reasoning provided by semantic technologies.

On the other hand, [15] proposes ORDM, an ontology based resource description model, where they describe resources in IoT environment. These resources are defined by the attribute, state, control, historical information, and privacy classes. However, the model does not provide flexible access control to captured data. Indeed, the users defined to access the IoT resources are fixed in the proposed ontology without any clear reasoning or criteria about their roles.

The Data Privacy Ontology (DPO) [2] defines its main class as IPEntity with three subclasses, namely DataHolder, Consumer, and Data which is linked with the PrivacyPolicy class. Although some privacy requirements are taken in consideration in the model, the solution mainly focuses on private data issues, and does not take in consideration industrial data. In addition, the OPD models some privacy policy terms, but does not consider IoT description concepts such as device, sensor, or observation.

In addition to SSN, Prov-O [8] is also recommended by the World Wide Web Consortium (W3C), and expresses the

PROV Data Model [14] using the OWL2 Web Ontology Language. It provides concepts to be used to represent provenance information generated in different systems and under different contexts. It can also be specialized to create new classes and properties to model provenance information for different applications and domains.

PROV Data Model is a model based on five requirements that must be met to allow provenance composition. First, a simple solution is constructed by which a message identifier is shared by its sender and receiver. The simple solution is then progressively fortified with bundles and topicIn property, attribution, and the pingback mechanism to satisfy traceability requirements. Although the solution proposed fills existing data traceability gaps, it still leaves open multiple issues, including system interoperability, and IoT environment description.

To sum up, it can be said that the existing solutions concerning ontology-based for preserving privacy in IoT did not address all types of data and focus mainly on private and sensitive data. Moreover, they focus only on who can access the data sensed by smart devices, and not on how it should be used. Despite the progress made by the discussed solutions, it seems necessary to propose a novel approach that rather than focus on security mechanisms such as cryptography, should cover data management issues by involving the different considerations cited in section 2.2. Furthermore, IoT devices must become self-sufficient by granting them the capacity to infer usage rights based on the owner's information security preferences. Thus, we propose IdSM ontology for IoT, data provenance and usage control, that provides a solution for overcoming the existing model limitations.

## 8 Conclusion and future directions

Internet of Things (IoT) is one of the key technologies in the industry 4.0 era and promotes the interconnection of numerous data sources in several sectors. However, the exploitation of data generated by the IoT resources raises security risks due to lack of trust. Moreover, the data owner loses complete ownership over his asset once it's being shared. Thus, semantic modeling becomes fundamental to infer the required information security to preserve the good functioning of the environment. For these reasons, we have proposed a new data sharing management ontology called IdSM-O that aims at defining a common vocabulary combined with standard reasoning technologies based on description logic and usage control to address information security concerns in the IoT environment.

Our research is a work in progress, and we are currently engaged in the technical foundation of a data provenance system as a security verification mechanism. Thus, when a data provider shares his observation output with its usage license, he can monitor from his own gateway whether the usage conditions he has defined are actually being respected



by data consumers. This research work is in line with several large schemes such as legislative procedures, audit, and quality management.

## Acknowledgments

This work is supported by the Conseil Départemental des Landes (PhD grant to N.Laamech).

## References

- [1] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. DBpedia: A Nucleus for a Web of Open Data. In *The Semantic Web*. Springer Berlin Heidelberg, Berlin, Heidelberg, 722–735. <https://doi.org/10.1017/S0034670500018052>
- [2] Narmeen Bawany and Zubair Shaikh. 2017. Data Privacy Ontology for Ubiquitous Computing. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 8 (02 2017). <https://doi.org/10.14569/IJACSA.2017.080120>
- [3] European Commission. 2020. *Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act)*. Vol. 55. COM/2020/767 final. Brussels, Belgium:.
- [4] Louis Dupré. 1993. The Common Good and the Open Society. *The Review of Politics* 55, 4 (1993), 687–712. <https://doi.org/10.1017/S0034670500018052>
- [5] Cuppens Frédéric and Miège Alexandre. 2004. Or-BAC, Organization Based Access Control. *The Review of Politics, Journées Druide* (2004).
- [6] Armin Haller, Krzysztof Janowicz, Simon Cox, Danh Phuoc, Kerry Taylor, and Maxime Lefrançois. 2017. *Semantic Sensor Network Ontology*.
- [7] I. Horrocks, P.F. Patel-Schneider, H. Boley, Said Tabet, Benjamin Grosz, and Mike Dean. 2007. SWRL: A semantic web rule language combining oWL and ruleML. *W3C Member submission* 21 (01 2007).
- [8] Krzysztof Janowicz, Armin Haller, Simon J.D. Cox, Danh Le Phuoc, and Maxime Lefrançois. 2019. SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics* 56 (May 2019), 1–10. <https://doi.org/10.1016/j.websem.2018.06.003>
- [9] Anas Kalam, R.E. Baida, P. Balbiani, S. Benferhat, Frédéric Cuppens, Yves Deswarte, A. Mieke, Claire Saurel, and G. Trouessin. 2003. Organization based access control. 120 – 131. <https://doi.org/10.1109/POLICY.2003.1206966>
- [10] Anja Klein, Hong-Hai Do, Gregor Hackenbroich, M. Kamstedt, and Wolfgang Lehner. 2007. Representing Data Quality for Streaming and Static Data. 3 – 10. <https://doi.org/10.1109/ICDEW.2007.4400967>
- [11] Timothy Lebo, Satya Sahoo, Deborah Mcguinness, Khalid Belhajjame, James Cheney, David Corsar, Daniel Garijo, Stian Soiland-Reyes, Stephan Zednik, and Jun Zhao. 2013. *PROV-O: The PROV Ontology*.
- [12] Jorge López de Vergara Méndez, Enrique Vázquez, Martin Antony, Dubus Samuel, and Lepareux Marie-Noëlle. 2009. Use of Ontologies for the Definition of Alerts and Policies in a Network Security Platform. *Journal of Networks* 4 (10 2009). <https://doi.org/10.4304/jnw.4.8.720-733>
- [13] Alexandre Miège. 2005. Définition d'un environnement formel d'expression de politiques de sécurité. Modèle Or-BAC et extensions. (01 2005).
- [14] Luc Moreau and Paolo Missier. 2012. PROV-DM: The PROV Data Model. (01 2012).
- [15] Shulong Wang, Yibin Hou, Fang Gao, and Songsong Ma. 2016. Ontology-Based Resource Description Model for Internet of Things. (2016), 105–108. <https://doi.org/10.1109/CyberC.2016.29>