Evaluation of Semantic Web Ontologies for Privacy Modelling in Smart Home Environments

Suzana Iacob[1], Antonis Bikakis[2],

[1] School of Management, University College London
[2] Department of Information Studies, University College London

[1] suzana.iacob.12@ucl.ac.uk, [2] a.bikakis@ucl.ac.uk

Abstract. The proliferation of smart devices gives rise to a new world of Ambient Intelligence, a world of technologies embedded in the surrounding environments, such as the home environment. As the success of such systems often depends on the collection on personal data, privacy concerns threaten to hinder this new world from reaching its full potential. At the same time, accurately modelling the different types of contextual information proves to be of paramount importance in paving the way towards the maturity of Ambient Intelligence systems, with Semantic Web ontologies becoming a popular solution. This paper aims to explore the application of Semantic Web ontologies in modelling privacy-related information in the context of smart home environments. To this purpose, we have conducted a practical evaluation of three ontologies, in an attempt to determine their suitability within the stated domain. The paper concludes that the representation of privacy features within smart home environments is attainable through the use of ontologies; however, current models do not achieve sufficient coverage of the domain. Lastly, the paper provides insights into practical ways of enhancing future ontologies in order to reach the required capabilities.

1. Introduction

Recent technological advances have enabled various devices with different capabilities to become embedded in our surrounding environments [1]. The ability of computer systems to seamlessly integrate into the lives of everyday users has been referred to by the term *Ambient Intelligence* (AmI) [2].

Given the sheer variety of connected devices that constitute an AmI system, privacy has become of paramount importance. This paper will explore the topic of privacy in smart home environments, from an information management perspective. In the context of this paper, information management broadly refers to the control, processing and exchange of information within a system.

A promising approach to information management for representing AmI domains is the use of Semantic Web ontologies [3]. These are tools that offer a shared understanding of a domain, enabling users to semantically represent relevant concepts from the domain.

The main objective of this study is to evaluate relevant Semantic Web ontologies in order to determine their degree of suitability for modelling privacy-related information in smart home environments. In order to meet the stated objective, the following research questions will be used to guide this study:

Q1. To what extent can current Semantic Web ontologies be used to model privacy aspects in smart home environments?

To address this question we have performed a task-based evaluation of three selected ontologies. The performance of the ontologies is assessed based on their ability to model real-life scenarios. The results of the evaluation are derived in two steps; first by determining the number of privacy features modelled; thereafter by validating against pre-determined evaluation criteria.

Q2. Are current approaches satisfactory? If not, what are their limitations and opportunities for future enhancements?

To answer this question, we identify the gaps and limitations of current solutions and point towards the requirements that future ontologies should meet and how current models can be extended to achieve these capabilities.

The rest of the paper is structured as follows: We first review the relevant literature to identify the main privacy challenges and privacy protection techniques in the context of smart home environments. The purpose of this exercise is to establish the key aspects of information that should be semantically represented in a system for smart home environments. In Section 3, we review current Semantic Web ontologies for Ambient Intelligence. In Section 4, we present the setup and the results of a task-based evaluation of such ontologies. Finally, in section 5, we discuss the main findings of the evaluation, identifying the main gaps and limitations of current approaches, and proposing guidelines and directions for addressing such limitations.

2. Privacy in Smart Home Environments

Ambient Intelligence (AmI) refers to systems that are adaptive, sensitive, and responsive to the presence of people [4,5]. Internet of Things is a new type of network through which everyday objects can communicate and exchange information. It can be viewed as an enabling technology for Ambient Intelligence. Within the AmI domain, a smart home is a residence containing "*ambient intelligence and automatic control, which allow it to respond to the behaviour of residents and provide them with various facilities*" [6]. Privacy, on the other hand, can be defined as the right of an individual to "*control the ways in which personal information is obtained, processed, distributed, shared, and used by any other entity*" [7]. The notion of privacy can be divided into hard and soft privacy. Hard privacy refers to practices which limit the amount of data shared, whereas soft privacy recognises the need to share data with other entities and instead employs techniques which control the conditions under which the data is being used [8]. Privacy concerns in AmI arose as early as the first AmI applications. Numerous studies in the literature mention various privacy challenges encountered in a smart home environment. By undertaking an analysis of the relevant literature, we summarize the primary privacy challenges in Table 1.

Table 1: Privacy Challenges

Challenge	**Definition**
Identity disclosure	Revealing information that can be used to uniquely identify a person; failing to ensure anonymity [9,10]
Sensitive information disclosure	Revealing data of sensitive nature such as biometric, health-related or financial data [4]
Personal data collection	Collecting data of a personal nature, in different ways and over a period of time [4,11]
Personal data centralisation	Storing and querying data from a central location [4]
Activity monitoring	Detecting and tracking human activity, typically using sensor technology [6]
Profiling	Constructing individual profiles based on data collected over time [2,4]
Personalisation	System performing actions to meet users' individual requirements [4]
Location Disclosure	Revealing data about an individual's geographic location [10]
Surveillance	Observing a person's activities over a period of time [2,6]
Unauthorised actions	System performing automated actions without the explicit consent of users [10]
Adaptability	Adapting to users' needs, typically by learning and improving over time [5,12]
Anticipation	Predicting the needs of users and acting accordingly [12]
Divergent privacy requirements	Conflicts of interest arising between the privacy needs of distinct users [9]
Personal data matching	Matching personal data from different sources in order to uniquely identify a person [10]

Multiple privacy protection techniques have been proposed in order to tackle the aforementioned challenges, as summarised in table 2. The majority of these are soft privacy measures such as purpose control and policies, while adequate security is a hard privacy measure.

Table 2: Privacy protection techniques

Protection Technique	**Definition**
Privacy policies	Specifying rules regarding data collection and sharing [3]
Authentication	Verifying the identity of a user or system [9,11]
Authorisation	Controlling who has access to what resources [10,11]
Adequate security	Preventing the potential misuse of information [2,13]
Purpose control	Ensuring that data is only used for its intended purpose [2,10]
Anonymizing Personal Data	Ensuring that data is not intelligible to other users other than the intended recipients [10,13]

3 Semantic Web Ontologies for AmI

Semantic Web ontologies meet the representation requirements of AmI set by many studies in terms of type and level of formality, knowledge sharing, expressiveness, flexibility and extensibility, generality, granularity, reasoning support and valid context constraining [14-16]. Context in this domain can be defined as "*any information that can be used to characterize the situation of entities [...] relevant to the interaction between a user and an application*" [17].

Several ontologies have been specifically developed for AmI systems [18-23]. After examining the range of available options, we decided to focus on SOUPA, COSE, and PROACT, as they originate from different domains, they therefore implement different modelling approaches and focus on different aspects of AmI systems, and they were all freely available to download.

SOUPA (*Standard Ontology for Ubiquitous and Pervasive Applications*) is a widely cited ontology, created specifically for AmI environments [18]. It is modular, and models the primary AmI features including *intelligent agents*, *space*, *time*, *events* and *policies*.

COSE (*Casas Ontology for Smart Environments*) is a relatively recent ontology, which achieves an in-depth representation of context features in smart environments, being highly domain-specific [20]. The main concepts represented in COSE are *sensor*, *building*, *occupant* and *human activity*.

PROACT (*PRivacy Ontology for ACTivity spheres*) has been specifically designed to reconcile privacy and AmI environments, being built upon concepts from general ontologies from the two fields, including SOUPA and Rei [23]. The key privacy features modelled are *resource*, *policy*, *policy mechanism*, *user* and *data processing*.

4. Evaluation of SW Ontologies for Privacy Modelling in AmI

The purpose of the evaluation is to draw conclusions regarding the suitability of the ontologies for modelling privacy in smart home environments The evaluation was conducted in three steps: We first analysed real-life scenarios, based on the privacy themes previously revealed by the literature review. We then attempted to model the scenarios using the selected ontologies. And finally, we assessed the capabilities of the ontologies to model the scenarios using appropriate evaluation criteria.

4.1 Scenario Analysis

Scenarios were created since the very first publications on AmI, starting with Weiser's portrayal of "Sal's World" [24]. The narrative depicts Sal and her daily activities, while devices anticipate her needs. Subsequently, other researchers developed scenarios to represent their vision of the future. Some researchers believed that AmI was portrayed as too "sunny", and developed a set of "dark scenarios" intended to raise awareness of the potentially harmful consequences, especially privacy threats [2]. The scenario "A Typical Family in Different Environments – Scene 1 (at home and at work)" shows a family, the Sebastianis, as they encounter challenges in their AmI home. Contemporary scenarios include modern elements in their view of AmI, such as social media and gamification. Denti [12] describes the "Butlers" as intelligent agents who aid the users while also having the ability to "entertain and make things nice". His scenario "Paolo & Archie" is an example of such functionalities.

Following the analysis of multiple scenarios, three were selected, as they contained most privacy elements: "Sal's World" [24], "Paolo & Archie" [12] and "A Typical Family in Different Environments – Scene 1" [2]; these will be referred to as Scenario A, B, and C, respectively, for convenience. The aim of scenario analysis was to explore the scenario text and identify the presence of privacy-related topics, along with the context in which they are found. By undertaking the literature review a set of privacy-related topics were revealed which were then mapped to the real-life scenarios. An example of the mapping is depicted in Table 3.

4.2 Evaluation of Privacy Elements

The next step was to attempt to model the features revealed in the analysis of the scenarios using the three ontologies. The experiment was carried out using Protégé, a widespread ontology development tool. Specifically, we created appropriate individuals and statements using the properties provided by the ontologies; an example statement from Scenario C is "*Paul owns PaulComputer*".

Through this exercise, the ontologies were examined from three different angles:

1. The extent to which they are capable of modelling the privacy features
2. Their ability to accurately portray a smart home environment
3. Their performance against the pre-determined evaluation criteria

In general, only small changes were made to the ontologies, such as adding appropriate subclasses to existing classes. If the required classes or properties did not exist, we concluded that the ontology failed to model that respective element. A sample of the findings along with the scenario mapping is presented in Table 3.

Table 3: Scenario-ontology analysis[1]

Scenario Excerpt	Privacy Elements	SOUPA Modelling
"Paolo's butler, Archie, detects that Paolo is coming home earlier, makes a tailored guess about the possible reasons, verifies the guess via personal messaging" (Scenario B)	*Activity monitoring*	✓
	Location Disclosure	✓
	Anticipation	~
	Unauthorised actions	✓

Some important findings of this experiment regarding the capabilities of the ontologies to model privacy-related concepts are the following:

SOUPA supports the representation of user privacy, being influenced by the Rei policy language [18], as can be noted from the class *Policy*. Privacy is protected by reasoning about the credibility of statements (e.g. *FalseStatement*, *reliabilityRating*) and about conflicting information (e.g. *conflicts*). SOUPA can model *Authorisation* through the use of *PermittedAction* and *ForbiddenAction*. However, we noticed that SOUPA cannot model authentication or the sensitive nature of data. In Scenario C, Paul's credit card information is sensitive, yet SOUPA has no means of expressing it.

COSE does not capture any privacy protection mechanisms, since no classes, properties or individuals are explicitly related to privacy. Wemlinger and Holder [20] do not mention whether privacy was considered as part of the development or whether it was planned as a future enhancement of the ontology. In the practical implementation of COSE, the only way to model privacy policies is through the *Plan* class, although this is arguably not the intended meaning of the class.

Privacy concerns are at the core of the PROACT ontology, as it was developed specifically to address this gap in modelling AmI systems [23]. PROACT has a *Policy* class, which can be used to specify user preferences and requirements. *PolicyMechanism* is the class used to propagate privacy protection throughout the system, having subclasses such as *Authentication* and *Authorisation*. A policy mechanism is added for a particular data-related action (class *Mechanism*) that a user or system can perform (e.g. *Authorisation* for *DataAccess*). PROACT displays the highest degree of privacy enforcing mechanisms. However, PROACT cannot model the prohibition of an action, such as the situation in Scenario C when incoming messages were blocked while the user was in a meeting.

The overall results of this evaluation are depicted in Table 4.

[1] Key to Tables 3 and 4:
✓/✕ = the ontology can/cannot model this feature
~ = the ontology can model this feature with minimum additions

Table 4: Privacy Evaluation Results[1]

Privacy Challenges	**SOUPA**	**COSE**	**PROACT**
Identity disclosure	✓	×	✓
Sensitive information disclosure	×	×	✓
Personal data collection	×	~	✓
Personal data centralisation	×	×	✓
Activity monitoring	✓	✓	×
Profiling	~	×	✓
Personalisation	~	~	~
Location disclosure	✓	✓	~
Surveillance	✓	✓	~
Unauthorised actions	✓	~	✓
Adaptability	✓	×	×
Anticipation	~	×	×
Divergent requirements	✓	~	×
Personal data matching	×	×	×
Privacy Protection Techniques			
Privacy policies	✓	×	✓
Authentication	×	×	✓
Authorisation	✓	×	✓
Adequate security	✓	×	×
Purpose control	✓	~	×
Anonymizing personal data	×	×	✓
Percentage of features captured by the ontology	**62.5%**	**27.5%**	**62.5%**

4.3 Evaluation of Ambient Intelligence Features

Previous studies on ontology evaluation suggest various possible evaluation criteria. In order to fully assess ontologies by their ability to represent facets of privacy, we carried out an analysis of general AmI features, since privacy is not a standalone concept within the smart home domain. Based on this analysis, we selected the evaluation criteria presented in Table 5.

Table 5: AmI features

AmI Features	Description
Context	Ability to capture, process, and explore the context in which the system-user interaction takes place [3]. Key features are *location*, *person*, *time* and *activity*.
Uncertainty	Ability to deal with data that is incorrect, imprecise, conflicting or incomplete [3]
Human-Computer Interaction (HCI)	Ability to model the interactions between users and the AmI system [3]

Thereafter, we evaluated each ontology and assigned performance scores ranging from 1 to 5. A score of 1 signifies a low performance in the respective area, meaning the ontology lacks the ability to model the required elements. On the other hand, a score of 5 represents a very high performance, suggesting the ontology accurately models the relevant AmI features.

4.3.1 SOUPA

Context **Score: 3/5**

The *Location* class is part of SOUPA, yet location data does not seem to be integrated well with smart devices. There were cases presented in the scenarios where a device collects location data, and this situation could not be modelled. There is only one *Person* class, and users are not differentiated from each other. The *Time* ontology contains numerous useful concepts such as instant events, events of longer duration and measurements of time.

Yet, the Activity ontology offers specific properties such as actor, target, instrument and time, enabling the formation of links between user activities and virtually any other part of the system.

Uncertainty **Score: 5/5**

SOUPA is the only of the three ontologies that is capable of modelling incomplete or uncertain information, with one of the core SOUPA ontologies being *Belief - Desire - Intention*. Using this ontology, we could create statements such as "*SmartHome believes IntruderStatement*", capturing a situation in Scenario C, where the system has information regarding an intruder in the home.

Knowledge is another key class for modelling uncertainty, having a *reliabilityRating* property, which can express the level of confidence regarding the validity of a particular piece of knowledge. Similarly, a piece of knowledge can conflict with, or be *inconsistentWith* another piece of knowledge.

HCI **Score: 4/5**

In SOUPA, HCI can be modelled through the use of the *Action* ontology, along with its specific properties (e.g. *actor*, *target*, *recipient*). HCI-specific means of communication are also captured, in particular, tools for messaging such as *ChatID* and *IMProvider*. These were employed for representing Scenario B, where Paolo and the intelligent system Archie were exchanging text messages. Yet the limited *Device* vocabulary restricts the richness with which HCI can be modelled.

4.3.2 COSE

Context **Score: 5/5**

COSE can model elements of surrounding space to a very high degree of accuracy. *Location* is captured through the *Point* class, having properties for representing coordinates (e.g. *xCoordinateOfPoint*). Space within a house is represented by the classes *Bedroom, LivingRoom*, and *FurniturePiece*. Moreover, COSE offers a *Person* class with the subclasses *Resident* and *Occupant*. *Time* is captured by the classes *TemporalThing* and *UnitOfTime* and through data properties such as *timestamp***.** *Activities* are classified into *IntelligentAgentActivity* and *HumanActivity*, a separation that we found extremely useful. Likewise, human activities are further broken down, achieving a higher degree of accuracy.

In addition, COSE distinguishes between the notions of *SmartEnvironment*, *ElectronicDevice* and *HouseholdAppliance***.** This approach facilitated the modelling of Scenario B, in which the intelligent agent Archie communicated with Paolo via an electronic device (i.e. phone) and scheduled the activities of household appliances (e.g. washing machine). Finally, COSE uses a highly precise hierarchy for representing sensors and actuators, an essential component of smart homes. The *Sensor* hierarchy is comprised of 14 subclasses including *PressureSensor*, *TemperatureSensor*, *MotionDetector* and *ContactSensor*.

Uncertainty **Score: 2/5**

COSE cannot explicitly model uncertainty. The only element that points towards uncertainty is the data property *activityHasError*, which can represent the uncertain nature of human activities; however its precise use is unclear.

HCI **Score: 3/5**

HCI modelling can be achieved only to a certain extent. COSE can model human activities and people's interactions with devices via the *InteractWithInformation* and *InteractWithPhysicalObject* classes. Nevertheless, the aspects modelled are incomplete; primarily due to the missing properties connecting the classes *Person* and *ElectronicDevice*.

4.3.3 PROACT

We should note that PROACT was not available in its entirety and therefore it was reconstructed by following the information available in [22]. Hence, due to lack of access to the original ontology, some aspects could not be fully evaluated.

Context **Score: 2/5**

Due to its focus on privacy and security, PROACT cannot model context sufficiently. The most noticeable omission is the lack of a representation for *location*. On the other hand, humans are well modelled by classes such as *User*, *Client* and *ResourceOwner*. In addition, PROACT allows the representation of groups of people as well as individuals, enabling the description of companies or other entities.

There is one class for *Time* and a property, *hasDuration*, but these only partially meet the time representation requirements of smart environments. In Scenario A, Sal was verifying the time markers of events in the surveillance system, elements which we could not model with PROACT. Finally, this ontology has no way of accounting for user *activities*, meaning all related aspects such as activity monitoring or surveillance, cannot be modelled.

Uncertainty **Score: 2/5**

PROACT does not have the ability to capture uncertain information. However, some level of uncertainty can be modelled as it relates to privacy and security. For example, user groups can collectively be assigned a "trust level", therefore, the ontology enables reasoning about the trustworthiness of users.

HCI **Score: 3/5**

HCI is generally well accounted for in PROACT. There are properties that enable stating that a user owns a device, and also that a device recognises the user. Through the class *Service*, intelligent agents can provide services that are received by users. This can be thought of as a class synonymous to "*Action*", yet only devices can execute these actions. Nonetheless, PROACT's capabilities to represent more complex HCI features, such as the exchange of instant messages mentioned in Scenario B, are rather limited, as it can only capture the type of a service, but not how this service is actually used.

4.4 Evaluation of General Quality Aspects

Aside from the specific AmI features, we also evaluated the ontologies using general ontology quality criteria selected from the literature, as outlined in Table 6. This evaluation was performed during the practical experiment by using the ontologies and deciding how well each ontology performed against the specified criteria.

Table 6: Quality Criteria

Quality Criteria	**Description**
Accuracy	The descriptions and definitions of terms are correct within the specified domain, capture the intended meaning, from the viewpoint of the users of the system [25]
Clarity	A term can be uniquely identified and distinguished from other terms, implying sufficient documentation and labelling [3,25]
Consistency	A concept should be defined in a coherent way without allowing for conflicts or contradictions and the ontology as a whole must be logically correct.[3,25]
Conciseness	The ontology does not contain irrelevant or redundant terms [25]
Completeness	The degree to which the domain is covered [25]
Operability	The degree of learnability, ease of use and memorability with respect to the ontology user [26]

The level of accuracy was given by the ability of an ontology to accurately model the concepts from the scenarios with appropriate classes and properties that associate those classes. An example of inaccurate class naming in PROACT is in the classes *Mechanism* and *PolicyMechanism*, which are not similar in function yet have similar names (*Mechanism* refers to an action such as transferring data, whereas *PolicyMechanism* could be authentication or authorisation).

Clarity was mainly derived from the presence or absence of clear naming, class hierarchies and labelling. For instance, COSE was evaluated as a generally unclear ontology, having unintuitive class names (e.g. *SupposedToBeMicrotheory*) and hierarchies (e.g. *Person* as a subclass of *ThreeDimantionalGeometricThing*).

Consistency was determined mainly based on whether there are any unsatisfiable classes or restrictions in the ontology.

The level of conciseness was given by the number of relevant and redundant classes. For example, COSE contains 145 classes, the majority of which have one subclass only; thus, they can be considered redundant.

An ontology was considered complete if during the scenario mapping, all needed elements could be found. For instance, SOUPA is more general-purpose and thus had no means of representing sensors, while PROACT lacked elements to represent contextual concepts such as location, time, space, and user activities.

Finally, operability was determined based on general conclusions about the ease of use of each ontology. The results of this evaluation are summarised in Table 7:

Table 7: Evaluation of Quality Aspects (All scores out of 5)

Evaluation of Quality Aspects	SOUPA	COSE	PROACT
Accuracy	3	3	3
Clarity	4	2	3
Consistency	4	3	5
Conciseness	4	1	5
Completeness	2	3	2
Operability	4	1	4
Average Score	**3.5**	**2.17**	**3.67**

5. Discussion and Recommendations

5.1 Results of Ontology Evaluation

The initial evaluation of the three ontologies determined their suitability in semantically representing privacy-related features in smart home environments. This exercise concludes that both SOUPA and PROACT could model 62.5% of the features, while, COSE is significantly behind, being able to model only 27.5%. Some features, such as *surveillance*, *location disclosure* and *unauthorised actions*, were representable in all ontologies, while *personal data matching* could not be modelled by any of the three ontologies.

The distribution of modelled features was homogenous; there were few features that could not be represented by any of the ontologies. It can, therefore, be concluded that it is possible to semantically represent all of the evaluated features, and the lower scores account for the missing capabilities of each ontology.

Nevertheless, it can be argued that an ontology which accurately represents privacy protection mechanisms is superior to one that can only represent privacy challenges. Consequently, the results were further analysed by considering the two facets of privacy independently, as shown in Table 8. This analysis differs from the previous one in the sense that privacy challenges and protection techniques are given equal weights. Previously, the challenges weighted more due to being more numerous (14) compared to the protection techniques (6).

Table 8: Privacy challenges versus protection techniques

Facets of privacy	**SOUPA**	**COSE**	**PROACT**
Privacy challenges	60.71%	35.71%	53.57%
Privacy protection techniques	66.67%	8.33%	83.33%
Average Score	**63.69%**	**22.02%**	**68.45%**

Therefore, PROACT is superior in modelling privacy protection techniques.

After the implementation of the ontologies in Protégé, conclusions could be drawn regarding their ability to model the principal AmI features. The results of the second evaluation are summarised in Table 9.

Table 9: Evaluation of AmI features (All scores out of 5)

Evaluation of AmI Features	**SOUPA**	**COSE**	**PROACT**
Context	3	5	2
Uncertainty	5	2	2
HCI	4	3	3
Average Score	**4**	**3.33**	**2.33**

Consequently, SOUPA performed exceedingly well in this evaluation with an average score of 4 out of 5, followed by COSE and lastly, PROACT. COSE was better in modelling context elements, yet SOUPA was the only one that could accurately represent uncertain and incomplete information.

Lastly, we assessed the general quality aspects of ontologies. In this assessment, PROACT scored the highest, followed by SOUPA and COSE. It is worth noting that these quality criteria are not fully independent to each other; for instance, an ontology that scored low in clarity, is unlikely to perform very well in operability.

For some criteria, accuracy and completeness in particular, neither of the ontologies scored more than 3 out of 5, meaning that their ability to describe context accurately is limited.

To conclude, based on the results of the three evaluation exercises, SOUPA and PROACT seem equally well equipped to model privacy in smart home environments, both having strengths and weaknesses alike. COSE has been deemed not fully suitable for this domain, requiring considerable enhancements.

Nevertheless, by observing the performance of the selected ontologies, broader themes emerge for the field of semantic web ontologies for privacy in smart home environments. The final section discusses these themes and offers an insight into what the future improvements in the field might be.

5.2 Conclusion on Ontology Evaluation

It is important to situate the findings within the wider context of how the ontologies themselves were developed. COSE is an ontology for smart environments, as Wemlinger and Holder argue in favour of developing domain ontologies [20]. Due to its narrow focus, COSE did not succeed in modelling privacy aspects. Thus, COSE has been perceived as an incomplete ontology, which should be enhanced with key themes in smart home environments, such as uncertainty and privacy.

PROACT is targeted at privacy and security, introducing highly valuable techniques for representing privacy protection. Yet, it disregards other context elements which, in turn, contribute to modelling privacy. The lesson learned from evaluating PROACT is that privacy must be considered holistically, accounting not just for privacy protection, but also acknowledging the privacy issues that arise from the very nature of smart environments. The most surprising finding is the failure of PROACT to outperform SOUPA. Not only is PROACT more recent than SOUPA, but also PROACT was built on top of SOUPA [23].

In spite of small shortcomings, SOUPA performed best overall, having no considerable gaps in modelling the domain. The justification could be that it is by design a general-purpose ontology, implying a careful consideration of the domain at large. The conclusion from this evaluation is that the most promising approach takes a holistic stance and regards privacy as an integral part of smart home environments.

Lastly, an interesting observation arises from the study of the relevant literature from different periods. The most recently developed ontology, COSE, as well as Denti's scenario [12] (Scenario B), paint a different picture compared to the visions in the early literature. Modern features, including precise representations of sensor technology, found in COSE, suggest that the visions of AmI are closer to being realised today, as the technological capability exists. Interestingly, Denti's scenario enhances the vision of smart homes with social media elements and gamification, suggesting that future ontologies should potentially be extended to capture the additional privacy challenges brought by these domains.

5.3 Recommendations

To begin with, the overarching theme that emerges from the evaluation is the fact that a suitable ontology must account for privacy protection as well as elements of the surrounding context. Consequently, future research could consider building on top of the context elements from COSE and the privacy protection mechanisms from PROACT. As a baseline overall structure, SOUPA stands out as being a promising starting point. In particular, the future ontology could benefit from a modular structure, similarly to the way SOUPA is organised, since the resulting ontology is likely to have a considerable size, which would reduce clarity and operability.

Thereafter, the ontology could be developed progressively starting from the core context elements: *Location*, *Person*, *Time*, and *Activity*. Building on top of these, the future ontology could integrate the COSE hierarchies for sensors and buildings, devices and household appliances.

Privacy protection measures should be added following the example from PROACT, by constructing classes for policy mechanisms such as authentication and authorisation, and the data-related actions that the policies will be applied upon (e.g. data access, data disclosure). In addition, we would recommend an extension that enables the permission as well as the prohibition of actions, so that the users can grant permissions both explicitly and implicitly.

Finally, the Belief-Desire-Intention construct from SOUPA has been deemed highly appropriate for modelling uncertainty. With regard to the properties, they should be structured in a hierarchical manner, grouped by common domain and range restrictions, as SOUPA proved this practice to be useful.

Nevertheless, future research should also bear in mind the limitations of ontology engineering, since automatically integrating ontologies is still an open research question [1]. Therefore, manual integration, or simply building a new ontology based on previous ones, could represent a viable option.

References

1. Augusto, J.C., Nakashima, H. and Aghajan, H. (2010). Ambient Intelligence and Smart Environments: A State of the Art. In: Nakashima, H., Aghajan, H. and Augusto, J.C. (eds.) *Handbook of Ambient Intelligence and Smart Environments*. New York, Springer US, pp. 3-31.
2. Ahonen, P. et al (2010). Dark scenarios. In: *Safeguards in a World of Ambient Intelligence.* Wright, D., Friedewald, M., Punie, Y., Gutwirth, S. and Vildjiounaite, E. (eds.) Dordrecht: Springer Netherlands, pp. 33-142.
3. Ye, J., Coyle, L., Dobson, S. and Nixon, P. (2007). Ontology-based models in pervasive computing systems. *The Knowledge Engineering Review*, 22(4), pp. 315-347.
4. Friedewald, M., Vildjiounaite, E., Punie, Y. and Wright, D. (2006). The brave new world of ambient intelligence: An analysis of scenarios regarding privacy, identity and security issues. *Security In Pervasive Computing.* 3934, pp. 119-133.
5. Cook, D., Augusto, J. C. and Jakkula, V. (2009). Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4), pp. 277-298.
6. De Silva, L., Morikawa, C. and Petra, I. (2012). State of the art of smart homes. *Engineering Applications of Artificial Intelligence*, 25(7), pp. 1313-1321.
7. Gritzalis, S. (2004). Enhancing Web privacy and anonymity in the digital era. *Information Management & Computer Security*, 12(3), pp. 255-287.
8. Deng, M. et al. (2011). A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering* 16(1). pp3-32
9. Wright, D., Gutwirth, S. and Friedewald, M. (2007). Shining light on the dark side of ambient intelligence. *Foresight*, 9(2), pp. 46-59.
10. Caire, P., Moawad, A., Efthymiou, V., Bikakis, A. and Le Traon, Y. (2016). Privacy Challenges in Ambient Intelligence Systems. Lessons Learned, Gaps and Perspectives from the AAL Domain and Applications. *Journal of Ambient Intelligence and Smart Environments*. In press.

11. Theoharidou, M., Tsalis, N., and Gritzalis, D. (2014). Smart home solutions for healthcare: privacy in ubiquitous computing infrastructures. *Handbook of smart homes, health care and well-being*.
12. Denti, E. (2014). *Novel pervasive scenarios for home management: the Butlers architecture.* SpringerPlus, 3(1), pp. 1-30.
13. Langheinrich, M. (2001). Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems. *Ubicomp 2001: Ubiquitous Computing*, pp. 273-291.
14. Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., Riboni, D. (2010). A survey of context modelling and reasoning techniques. *Pervasive and Mobile Computing* 6(2), pp. 161–180.
15. Bolchini, C., Curino, C.A., Quintarelli, E., Schreiber, F.A., Tanca, L. (2007). A data-oriented survey of context models. SIGMOD Rec. 36(4), pp. 19–26.
16. Li, Y., Pan, J., Krishnaswamy, S., Hauswirth, M. and Nguyen, H. (2014). The Ubiquitous Semantic Web: Promises, Progress and Challenges. *International Journal on Semantic Web and Information Systems*, 10(4), pp. 1-16.
17. Dey, A., Abowd, G. and Salber, D. (2001). A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications. *Human-Computer Interaction*, 16(2-4), pp. 97-166.
18. Chen, H., Perich, F., Finin, T. and Joshi, A. (2004). SOUPA: Standard ontology for ubiquitous and pervasive applications. *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pp. 258 - 267.
19. Wang, X., Zhang, D., Gu, T. and Pung, H. (2004). Ontology based context modeling and reasoning using OWL. *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*, pp. 18-22.
20. Wemlinger, Z. and Holder, L. (2011). The COSE Ontology: Bringing the Semantic Web to Smart Environments. *Toward Useful Services for Elderly and People with Disabilities*, 6719, pp. 205-209.
21. Wongpatikaseree, K., et al. (2012). Activity Recognition Using Context-Aware Infrastructure Ontology in Smart Home Domain. *2012 Seventh International Conference on Knowledge, Information and Creativity Support Systems*, pp. 50-57.
22. Coutaz, J., et al. (2003). *Working document on gloss ontology*. Technical Report D9.2, Global Smart Spaces.
23. Panagiotopoulos, I., et al. (2010). PROACT: An Ontology-Based Model of Privacy Policies in Ambient Intelligence Environments. *2010 14th Panhellenic Conference on Informatics*, pp. 124-129.
24. Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*, 265(3), pp. 94-104.
25. Vrandečić, D., (2009). *Ontology evaluation*. Springer Berlin Heidelberg. pp. 293-313.
26. Duque-Ramos, A., et al. (2011). OQuaRE: A SQuaRE-based Approach for Evaluating the Quality of Ontologies. *Journal Of Research And Practice In Information Technology*, 43(2), pp. 159-176.
27. Antoniou, G., Groth, P., Van Harmelen, F. and Hoekstra, R. (2012). *A Semantic Web primer*. 3rd ed. London: MIT Press.