Semantics for incident identification and resolution reports

JOAQUÍN BORREGO-DÍAZ, ANTONIA M. CHÁVEZ-GONZÁLEZ,
JOSÉ L. PRO-MARTÍN and VIRGINIA MATOS-ARANA

Department of Computer Science and Artificial Intelligence – University of Seville, Spain.

Abstract

In order to achieve a safe and systematic treatment of security protocols, organizations release a number of technical briefings describing how to detect and manage security incidents. A critical issue is that this document set may suffer from semantic deficiencies, mainly due to ambiguity or different granularity levels of description and analysis. An approach to face this problem is the use of semantic methodologies in order to provide better Knowledge Externalization from incident protocols management. In this article, we propose a method based on semantic techniques for both, analyzing and specifying (meta)security requirements on protocols used for solving security incidents. This would allow specialist getting better documentation on their intangible knowledge about them.

Keywords: Incidents, security, semantic technologies, knowledge management.

1 Introduction

A key organizational activity in Security for Information Systems (SIS) is the document generation, spreading and management, for both employees and clients (see e.g. Microsoft document [24]). The documents meet several purposes: spread information within and outside the organization, facilitate self-learning within the organization, make explicit tacit knowledge on incident management, isolate emergent security concepts, etc. Documentation service provides robust strategies and secure solving methods to face a wide range of situations. Among the documents, these related with reports on incidents, protocols and information on systems can play a structural role in the SIS paradigm. Their role is not limited to Document Engineering (DE) as it covers several levels, for example, documentation, diffusion and self-learning among employees.

However, as it is said in Mace et al. [19], reports generally describe information security policies by mixing professional opinion, staff experience, technology manufacturer advice and external security standards or regulations. Therefore, it is natural to think that traditional and successful methods and policies have to be documented in order to get better knowledge externalization (KE) in the sense of the classic framework introduced by Nonaka and Takeuchi [23]. KE represents a clever strategy in SIS documentation since reports are useful mainly for organization members (which share the same tacit knowledge about this), and it is frequent that new paradigms force them to conciliate knowledge.

This scenario of deficient KE contrasts with the current ubiquitous role of SIS, which has evolved from a technical discipline to a strategic concept (see e.g. US National Academic briefing [22], Smith and Spafford [34] and, in the cloud paradigm, Catteddu and Hogben [9]). The world's growing dependence on a powerful but vulnerable Internet—combined with the disruptive capabilities of cyberattackers—now threatens national and international security. Thus, it is even necessary to think on the problem from the point of view of Complex Systems Science (Sadvandi et al. [28]).

In Geer's book [13], influence factors in security incidents are summarized, showing the complexity and hardness of the problem. Author juxtaposes cyberattack advantages and mitigation strategies according to their level of influence of one to the other. He presents different features of attacks at general level, and also considers SIS vulnerabilities and defences classified in different categories. The study suggests that a clear and robust classification could provide a better defence position. Robust classifications have to be stable under comparisons with other approaches. Therefore, a scientific approach to cybersecurity is indispensable, even as a target for (e-Semantic) Science, covering seven interrelated themes (see Riley [32]): Common Language, Core Principles, Attack Analysis, Measurable Security, Risk, Agility and Human Factors. All of these themes converge in document representation in SIS.

The other challenge associated with SIS documentation is the potentially vast number of disparate information sources, which makes the management of SIS information complex and time-consuming (cf. also Mace et al. [19]). Although KE consolidates the knowledge within individual organizations, it is typically kept 'in-house' and the interoperability among different organizations could be a challenge. That is to say, it may suffer of interoperability issues, most of them of semantic nature. This gap makes harder to exploit SIS documentation of an organization by others or its effective spreading, a problem related with the global nature of SIS challenges.

1.1 Semantic methods for SIS

Semantic web technologies (SWT) can provide an unified point of view solving the aforementioned problems. It is important to point out that SWT can be viewed as both, a technology and a scientific discipline for knowledge representation and reasoning (KRR), which facilitates knowledge extraction, representation, management and even reasoning, instead of a discipline focused only on Ontology Engineering. The last viewpoint is very useful to face the above challenges by considering those as Knowledge Management Problems grounded on DE (cf. Glushko and McGrath [14]).

By focusing on the theme of this article, two related tasks matter. On the one hand, the attempt to formalize the information described in the reports provides knowledge emergence. On the other hand, SWT naturally solve interoperability problems by means of KE. Consensus efforts to represent document's knowledge by means of ontologies and data allow engineers and employees getting a sound understanding of ideas (which can be externalized), represented by means of concepts, properties and axioms of the ontology. Therefore, the problem of understanding the structure of concepts to anticipate potential issues of document information may be solved by the joint work of KRR specialists and security experts. Such solutions could be provided during activities driven to ontology creation, instead of only exploiting the final product, i.e. the ontology.

Additionally, SWT provide tools for analysing important features as, for instance, consistency, compliance with current Security Standards, as well as the fidelity with the intended model (see e.g. Aranda et al. [3] in SIS, and Alonso et al. [2] in general). The latter feature is based on a sound representation of some concepts, i.e. to say, whether the specification represents the intentions of security experts and whether there are axioms (or properties) clearly incompatible with real concepts. Therefore, the ontology-based approach enables the definition of security concepts and their dependencies in an understandable way for both humans and software agents (see Pereira et al. [27]).

1.2 Aim of the article

The aim of the article is to exploit SWT and associated methods, in the analysis and refinement of knowledge from security reports. The idea is to get better documentation by means of a semantic evaluation and improvement proposals. The idea bases the process of ontology creation on information contained in the documents.

The document set published by Spanish INTECO–CERT institution[1] has been selected as running example: *INTECO's identification and report of security incident for strategic operators* [36] and *The operator console. A Basic Guide to Critical Infrastructure Protection* [37]. The first one aims to be a guide intended to serve as a manual for action reporting and management related to Critical Infrastructure and Strategic Operators incidents through the INTECO–CERT. The second one describes the actions that operators have to perform in order to provide an effective and efficient response to security incidents. The documents provide a standardized protocol effort for both, effectively solving and documenting security incidents in a SIS scenario.

1.3 Structure of the article

The structure of the article is as follows. The next section is devoted to analyse relevant features in security documentation, with particular emphasis on the use of flowchart descriptions as main tool in processes description. In Section 3, the strategy we propose is described. In Section 4, some relevant results on the application of the strategy to a specific case (an Incident Report and Identification document) are reviewed. Section 5 provides some hints about the evaluation of the strategy within Knowledge Management Framework. Lastly, some conclusions on both the strategy and its applications are discussed (Section 6).

This article is an extended version of the work by Borrego et al. [8].

2 Semantic features of security documentation

The analysis of SIS documents has to be performed from different points of view, by distinguishing between classification of SIS elements (e.g. identification of incidents) and the description level of security protocols (for reporting or solving incidents). The representation of different features will provide essential elements (classes and particular individuals) for the ontology in a natural way. Due to the modular nature of the ontology, it should be allowed to extend or modify these elements without a general reconsideration of ontological commitments. To achieve this modularity, the top level of the ontology have to conciliate both points of view (identification and reporting), while low-level classes will represent a set of particular elements (usable actions, specific protocols, a set of possible identifications and classifications, etc.). Identification and protocol descriptions have different ontological nature although they share some common features allowing to articulate the ontology in two sub-hierarchies.

Although it would be possible to specify identification and resolution protocols by means of standard service ontologies (e.g. OWL-S or WSMO), a specific flowchart-based light ontological description of protocols is created instead. The reasons of this choice are justified by the particular features of SIS documents as well as by the details of protocol description within them:

- Description (at operator level) is simpler than that of standard service ontologies: a succinct and clear specification is better to understand protocols than a complete one. It is also more adequate for pragmatic representation of solutions.

[1] Acronym of Spanish Incident Response Center Security http://www.inteco.es/home/national_communications_technology_institute/

- The representation of protocols in documents is very similar to their natural (graphical) descriptions, making them easily understandable by operators and, in general, by any SIS organization member.
- A concise semantic description of the protocols which does not add complexity to reasoning services is provided.
- Because of natural mapping between actions and their corresponding ontological representation, the addition of new actions/description elements does not require SWT experts.

A secondary goal of the formalization of flowcharts is that the insertion of a new flowchart in the systems will be guided by the requisites flowchart ontology has (which can be understood as questions to be completed by the user). Since the flowchart ontology will be used, it is interesting to briefly describe it.

2.1 A view on the flowchart ontology

Flowcharts are the standard way to represent processes, or, in general, almost every kind of protocols involving dynamics in terms of state changes. Flowchart is designed as a directed graph in which nodes are represented by boxes and edges (arcs) are represented by arrows showing the process flow in the diagram.

The dynamic dimension of semantic analysis of SIS guides will be obtained by means of a flowchart-based representation of protocols. The version of the basic concept on the ontology is depicted in INTECO terms in Figure 1 from [36]. A singular feature of the ontology is the identification between `AtomicAction` and `FlowchartAction` classes. This non-orthodox equivalence is the result of a group discussion among authors. Ontological distinction between action and representation of the action within flowcharts is discarded. In this way action class is used in both levels, and non-experts in Ontology Engineering will understand the flowchart better.

There are multiple variants of flowcharts (Petri nets, ASM charts and so on). The simplest one only uses two types of nodes (boxes): *Action boxes* and *Decision boxes*. The first ones contain a set of actions that the user should execute in that state, therefore an action box must have one and only one output path. They are represented by the class `ActionBox` in our ontology.

The second kind of boxes are the *Decision* ones, where the inner text is a condition to be verified. The next current state depends on the value at which the condition may be evaluated (the condition is not a logic formular to be evaluated within the ontology, it is represented as a simple text). This kind of nodes can have multiple output paths. Decision boxes are modelled by the class `DecisionBox` in our ontology. Figure 1 shows the hierarchy of classes of the sub-ontology. It can be understood that `ActionBox` and `DecisionBox` are subclasses of a more generic concept, which we have called `InnerBox` (representing the internal nodes of a flowchart). In this way some restrictions on

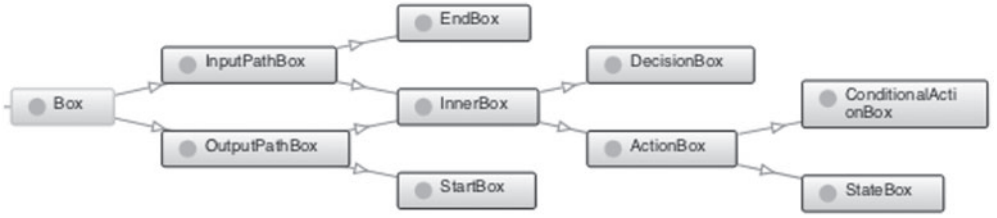

FIG. 1. Flowchart box element class.

the classes can be added, as for instance:

$\texttt{ActionBox} \sqsubseteq (= 1\ \texttt{hasOutputPath.Path})$, $\texttt{DecisionBox} \sqsubseteq (\geq 1\ \texttt{hasOutputPath.Path})$

Some types of flowcharts have two special nodes. Those that do not have an input path (i.e. input degree in the graph is equal to zero) and those that do not have an output path (i.e. ouput degree is zero). These nodes are represented in our flowchart ontology by means of `StartBox` and `EndBox` classes, respectively. We can enforce these constraints making these classes subtypes of `OutputPathBox` and `InputPathBox`:

$\texttt{InputPathBox} \sqsubseteq \exists\ \texttt{hasInputPath.Path}$, $\texttt{OutputPathBox} \sqsubseteq \exists\ \texttt{hasOutputPath.Path}$

Thus, an instance of `InnerBox` must inherit both restrictions:

$$\texttt{InnerBox} \sqsubseteq \exists\ \texttt{hasInputPath.Path},\ \texttt{InnerBox} \sqsubseteq \exists\ \texttt{hasOutputPath.Path}$$

Some other key concepts and classes of this ontology are `Condition` and `Path` with the natural associated semantics.

As it was already mentioned, there exists a number of ontologies that could be used for flowchart semantic representation. Among them, similar approaches to be presented here are related with the representation of industrial/business processes. Also the flowchart concept appears in biotechnology ontologies as, e.g. in the National Cancer Institute Thesaurus,[2] Semanticscience Integrated Ontology[3] and SWEET Governance Ontology[4], among others approaches.

3 A strategy for KE from incident reports

The goal is to enrich KE processes by means of Ontology Extraction processes. In this article, we propose a strategy for ontology extraction to accomplish this aim. The strategy consists of four stages (Figure 2):

Stage 1: Preliminary analysis: In this stage, activities are closely related with the rough understanding of goals and reports structure of the organization:

1. To state the scope and intended use of SIS document. A first distinction between description-oriented and solution-oriented protocols and methods is made.
2. Document analysis. Ontology engineers analyse the logical structure of the document and isolate main concepts used within it.
3. To determine the ontological nature of different concepts. Elaboration of a first categorization (possibly by building several hierarchies).
4. To find potential ambiguities or deficiencies in elements to be included in the ontology.

The Results expected in Stage 1 are related with the above analysis; formal and specific definitions are still not required. These are focused on the understanding of:

- Scope and Intended use.
- A high-level categorization, a first set of concepts extracted from document analysis and a first approach on the relationship structure among the different elements.

[2] http://bioportal.bioontology.org/ontologies/NCIT
[3] https://code.google.com/p/semanticscience/wiki/SIO
[4] http://wiki.esipfed.org/index.php/SWEET_Governance

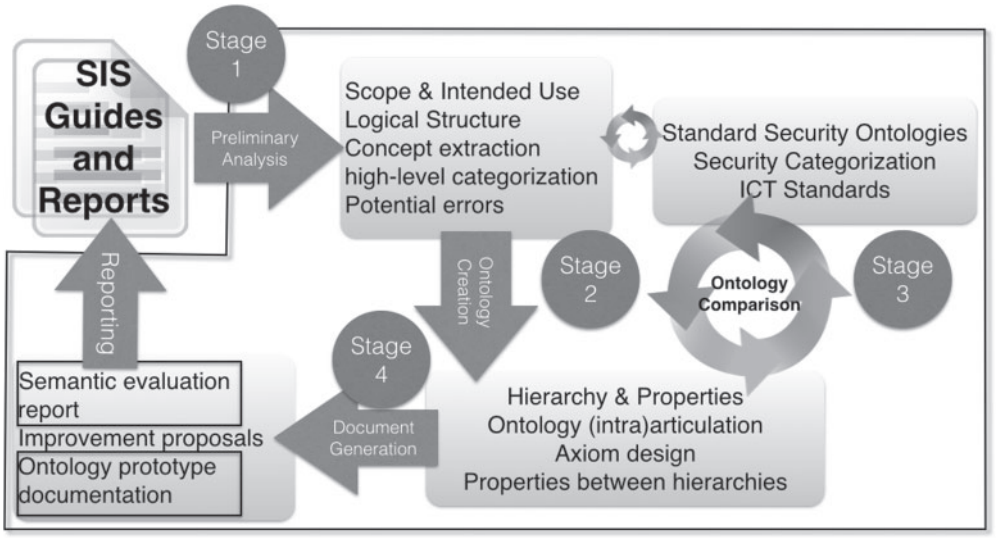

FIG. 2. Strategy applied to SIS documents.

- Above results may reveal potential errors. They are compiled to be solved later.

Stage 2: Ontology creation activities: The adoption of a pre-existent security ontology to formalize and clarify the SIS documentation, does not seem a sound approach, given the goals. The main reason is that it describes an approach to SIS report/classification that could be incompatible with the tacit knowledge (in particular intangible assets) of the organization. Since the main goal is KE instead of the production of a stable ontology, the activities to be performed in this stage are classic steps in ontology building:

- Hierarchies and properties implementation (e.g. using Protégé).
- Design of axioms (classes specification) for the key concepts.
- Study of relationships between sub-hierarchies devoted to different KRR problems (e.g. descriptions and methods).

Results expected in this stage are:

- An explicit representation of Hierarchy and Properties
- Axiom design (only natural axioms are added here), some of them devoted to describe properties between hierarchies

Stage 2 produces a first formal approach of tacit knowledge from SIS organization through the documentation. This approach allows SWT experts confronting this issue with other known representations (mainly ontologies).

Representability of security issues. The proposed bottom-up approach is the natural choice as it is not primarily intended to build a (other) security ontology. The aim is to build an explicit representation of tacit consensual knowledge (concepts, methods, relations) in reports documentation within an organization (this ontology may be a consensus on concepts and terms). Since bottom-up ontology building strategy is grounded on security information resources, and, since these kind of resources have not been designed to fit ontological structures, several deficiencies of representation

TABLE 1. Representational problems in Information Security

[P1:] No concepts for some kind of vulnerabilities	[P2:] Vague connections between threats and controls
[P3:] No relationships between threats	[P4:] Inconsistent granularity of information
[P5:] Redundancy and overlapping of information	

arise. In [11], Fenz and Ekelhart detect a number of representational problems when enriching a security ontology by means of Information Security (see Table 1). The bottom-up process aids to solve most of these problems (P1,P2,P4,P5), while problem P3 rests explicitly posed (to be solved by SIS experts). It is worthy to note that the adaptation of a general security ontology for this task is hard to automate, because some revision criteria cannot be fully formalized.

Stage 3: Comparison with standard security ontologies: The activities are driven to repair, refine and enrich the knowledge extracted in the above stage. The comparison strongly depends on a number of KRR issues. Likewise it is worth to note that the task has to be made both automatic (e.g. by merge-ontology tool of Protégé) and manually (by discussing other non-directed relationship among classes). In this case, manual comparison is essential because engineers attempt to redefine security concepts for standardizing emergent concept with previously established ones in other ontologies.

Results expected are encompassed in an analysis on the relationship between the prototype ontology and well-known Security Ontologies, Security Categorizations and in general other ICT Standards [6, 11, 16, 19, 26, 27, 30, 33].

Stage 4: Semantic evaluation report (with improvement proposals): Each step requires some discussion on features of key concepts. Such discussion has to be documented in the final report. The use of the ontology as a semantic reference of future SIS documents has to be taken into account.

Activities: SWT engineers have to document all the tasks and decisions of the strategy, concluding with a document on recommendations.

Results expected are oriented to the client (SIS organization), by documenting the strategy and results: a semantic evaluation report which includes improvement proposals. Finally, an ontology prototype documentation is provided, if experts consider that it is interesting to publish the ontology.

4 A case study: Applying the strategy to Incident Report and Identification documents

This section is devoted to summarize some observations about the selected stages of the strategy, as well as to discuss the main conclusions on the application of the presented strategy to (Incident Report and Identification) IRD documents [36, 37].

4.1 Phases of incident response. Some results from Stage 1

According to [36], the description of the main phases in incident response and mitigation of risk are (see Figure 1, from the document [36]): *Identification* (classification), *contention* and *mitigation*, *evidence preservation* and legal considerations, documentation and recovery. The elements of these phases have different nature. On the one hand, classification and identification have static nature while actions correspond to protocols (non-complex plans).

4.2 Static versus Dynamic dimension (from Stage 1, activity 2)

Preliminary analysis of documents (stage 1) shows that two particular ontological dimensions are combined. The first one refers to (statical) identification of main elements. This fact is important due to the fact that solving/repairing/mitigation methods strongly depend in the secure identification of the incident, which depends on turn of the classification of them. Despite that, it is hard to state the complex relationship among different categories. SIS documents often enumerate elements appearing in a particular organization, and the methods often depend on such classification. However, refinements of the categorization aid to specify the methods.

4.3 Ontology creation activities (from Stage 2)

Stage 2 includes the description of dynamic elements, e.g. protocols and methods. A protocol description is more precise than risk identification. This observation suggests building a flowchart-based sub-ontology in order to describe them. This ontology was described in Section 2.1. Likewise, an ontology on the incidents and processes described in the document was built. As it was already mentioned, such ontology represents a formal description, useful to compare intangible knowledge from the document with other well-established formal representations from the following stage.

4.3.1 Logical specification of meta-security in IRD (from Stage 2)

Specification of the ontology opens the possibility of including constraints that would be included in the SIS documentation (in natural language). Some of them would allow to monitor integrity/safety constraints. For example, the system only considers as *detected incident* the one for which it has an evidence:

$$\texttt{Detection} \equiv \exists \texttt{hasEvidence.Evidence}$$

Likewise, flowchart semantic specification allows instantiating protocols, making each one a complete and consistent representation of a security method. In particular, only flowcharts representing approved methods can be included:

$$\texttt{FlowChart} \sqsubseteq (\geq 1\ \texttt{represents.Action})$$

where $\texttt{Action} \equiv \texttt{AtomicAction} \sqcup \texttt{ProceduralAction}$. The absence of classification of an incident is prevented by a restriction axiom on the property `originIn`:

$$\texttt{Incident} \sqsubseteq (=1\ \texttt{originIn.Risk})$$

4.4 Analysing features by comparison with other ontologies (Stage 3)

This step includes evaluating the compatibility of implicit knowledge within INTECO–CERT document with other formalized approaches. The semantic description of SIS has a great advantage that allows comparing the INTECO–CERT approach to risks with other related classifications and/or ontologies, in order to evaluate its soundness—specifically by ontological mapping with other pre-existent risk hierarchies. This way analysis can be benefited from the comparison in order to: detect important absences in risks catalogue, isolate redundancies, explicitly reflect about particular risks as well as the possibility to propose other ontologies enrichment. It is worth to note that concept

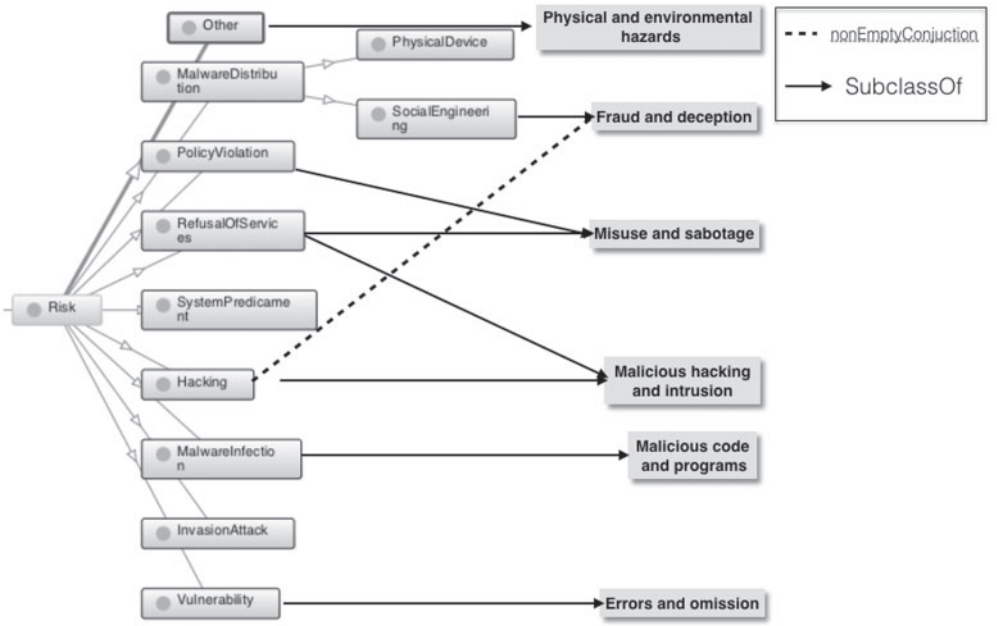

FIG. 3. `Risk` class and its relationship with categories from Smith et al. [33].

mapping between these general categories and the INTECO–CERT ones provides insights useful to enrich the description of action classes related with them. Those indications could be incorporated in the final documentation.

It is particularly interesting to consider its relationship with the following six general categories of information technology risk (according to Smith et al. [33]). The relationship among both categories is depicted in Figure 3. This relationship has to be understood as a set of incipient refinements of the ontology. Nevertheless, it is interesting to highlight some of them:

- *Malicious code and programs*: The concept contains `MalwareInfection`. Thus, the ontology could be expanded by adding classes to prevent risks. It requires protection at the individual and system level.
- *Malicious hacking and intrusion*: contains `Hacking` and `InvasionAttack`. However, INTECO classification also considers malicious hacking without intrusion (`RefusalOf Service`).
- *Fraud and deception*: According to Smith et al. [33] various forms of attacks in the form of spoofing, masquerading or salami attacks have been used to damage privacy. Common electronic forms of fraud include phishing and credit card theft. The first type corresponds to `SocialMalware` and part of `Hacking` while the second one correspond to `Social Engineering`. In this case the ontology is more specific than the category from [33]. In Figure 3, dashed line indicates that:

$$\texttt{Hacking} \sqcap \texttt{FraudAndDeception} \not\equiv \bot$$

- *Misuse and sabotage*: Closely related with `Vulnerability`. It also contains the class `Policy Violation`. The first class is one of the underspecified concepts in INTECO–CERT. The

TABLE 2. Viability analysis

Business viability	High. It is straitghforward to show advantages and the expected added value.
Technical viability	High. Technical documentation has a semistructured form, which facilitates the application of the strategy.
Project viability	Actions described in the stages are feasible and it is expected that organization teams accept reasoned changes and methods.

original category from [33] represents the resources that can be misused, or vandalized through unauthorized access. One form is unauthorized software changes.

- *Errors and omissions*: Closely related with `Vulnerability`. According to Smith [33], this category assumes accidental (software) errors, to include unintended destruction of files or data, as well as routing or transmission errors. This also includes programming errors. Thus, it seems that `Vulnerability` class has not a good granularity level in INTECO document.
- *Physical and environmental hazard*: It is out of the scope of the `Risk` class of [36].

The ontological analysis of this kind of relationships among categorizations can be also made in other parts of the ontology, by using other similar security ontologies. Even it can induce to distinguish between safety and security, in order to refine ontology in some SIS scenarios (Chapon et al. [29]).

A more detailed classification and description of risks needs the formal inclusion of *damage* concept. This inclusion would force to refine risk categories, as in Kim et al. [18]. Also, it is interesting to refine concepts about cyberattacks from Geers' [13]. In this way, the inclusion of *target* concept allows the introduction of new mitigation strategies at the dynamic level.

4.5 Dynamic ontological level

One of the INTECO–CERT tasks is the response to security incidents reported, as occurring in Critical Infrastructures, by users of the service, ensuring that the relevant information is stored. The description of the process follows the scheme shown in Figure 1 of [36], which can be fairly represented by means of the flowchart representation.

The stage of (internal) ontology articulation allows building semantic bridges among the sub-ontologies. In fact, descriptional and dynamic ontologies share concepts of common ontological nature. This step produces the refinement of the high level of the ontology.

4.6 On ontology population

With respect to ontology population, two main kinds of individuals for ontology population can be extracted from documents (protocols and incidents). Ontology tools provide interfaces to guide operators in the population of different classes. In fact, populating security ontology with methods and use cases also suggests the need of extending the information of the document. It is important to consider a description of detection mechanisms with the aim of detecting anomalies in the operating system itself, namely the HIDS (host-based intrusion detection systems) in [37]. The idea is that HIDS provides a complete incident description. This consideration induces to think that *Agent* contains humans and artificial agents.

TABLE 3. General features of the strategy

Facet	Feature	Evaluation
Influence of the application domain	Critical environment	High. Although it is a critical environment, conclusions come from ontology specifications that comes in turn of company's knowledge.
	Acceptability by end-users	High. Ontology has its origin in the specification of organizational culture on reports.
	Partial acceptability	High. Knowledge provided by ontology analysis is modular: It is possible to isolate partial recommendations if necessary.
	Each stage	High. Any phase interact with SIS systems, therefore only stage 4 can be considered as an influence phase.
Validation	As experimental task	Performed on historical cases (former security incidents)
	As a KE method	High. Once document recommendations are accepted, the product and documentation of the strategy rests integrated in the documentation ecosystem of the company.
Synthesis of new concepts	New concepts and procedures	Medium. Simulation (non-experimental) against SIS operators could be performed.

4.7 On results of Stage 4

The ontology specification opens the possibility of including constraints, which would be included in SIS documentation. Actions and indications proposed to the organization from the previous two stages are collected, and specific elements of action are recommended: in what areas reflection on the current documentation should focus, as well as which is the recommended final solution in each case, etc.

5 Evaluation

Since the process aims to debug and clarify security reports by means of their correct specifications (by using ontologies as tools), the evaluation of the method should be based on report authors' feedback (validation by the domain expert). A secondary product (the ontology containing formal specifications) is useful to compare reports knowledge with that represented by standard ontologies on both facets, the scope and the intended use (validation by the knowledge engineer): ontology soundness produced by means of this strategy is useful to revise the report itself. Thus, the ontology is, in this case, a tool instead of a goal. Although the use of ontologies built from standard security

TABLE 4. Evaluating the strategy as a KM process according to KM^2 initiatives

	Objectives	Qualitative evaluation of the proposed strategy (outside of the scope, low, medium, high)
Culture	KM is an integral part of the organisational culture	Outside of the scope, not considered
	KM enables collaboration between experienced and inexperienced personnel	High. It enables collaboration by means of improving documentation
	KM encourages and facilitates the exchange of organizational knowledge	High
Organization	KM defines the organisational structure	Outside of the scope
	KM supports interdepartmental collaboration	Regular. Strategy does not propose interdepartamental teams
	KM supports the collaboration between employees and managers	Low. Indirect support by means of stage 4 recommendations
Methods	KM practices are integrated into knowledge-intensive work processes	Low. Process is external to organization
	KM supports the integrative (synchronised) approach to managing implicit and explicit knowledge assets	Outside of the scope of the method
	KM supports the exploration, innovation, dissemination and automation of knowledge	Low. Indirect support. Medium if ontology is finished and published
Processes	KM supports the establishment of continuous business processes	Outside of the scope
	KM supports the reduction of work processing time	Low. Indirect, because describe better identification procedures
	KM supports the avoidance of work redundancy	Low. Indirect: better classification of security incident allows solving procedures

descriptions are very useful to enhance the behaviour of multiagent systems for security issues (see e.g. Herrero et al. [15]), its use is out of the topic of this article.

TABLE 5. Evaluating the strategy as a KA process

Features of KA extraction	Evaluation of the proposed strategy
Preprocessing requirement	Lesser requirements. Only the avaibility of technical reports on SIS
Ontology reusability	Low. Ontology is a secondary result and it is very specific
Extraction level	High. Due to the qualitiy of technical reports
Degree of automation	Low. In fact only in Ontology comparison automated tools assist to analysts in Stage 3
Algorithm selection	Not considered
Efficiency	Not considered
Reliability	Only changes on intern knowledge of tech reports have to be evaluated against domain experts

The abovementioned features do not represent the overall framework to evaluate the strategy: its evaluation of the proposed strategy can be considered from other important aspects. The strategy is, in fact, a knowledge management (KM) one. Also the strategy can be viewed as a knowledge acquisition (KA) method, thus it could be evaluated in this classic framework (cf. Shadbolt et al. [31]). Therefore, evaluation is a risky and complex issue.

The aforementioned observations lead us to consider three features. The first and third ones are related to its KA nature whilst the second one considers its KM nature.

Viability analysis. While results are evaluated by using the documents generated in Stage 4, it is also interesting to consider whether the application of the strategy has high impact on the organization. See Table 2 where a summary on three types of viability analysis is shown.

Evaluating the strategy itself. This facet addresses to its nature as KM Method. Table 3 collects the main conclusions on this point of view. From the point of view of KM processes, qualitative evaluation according to the prospective monitoring process KM^2 is described in Table 4 (see Minonne and Turner [21]). From the point of the view of KA (by considering the strategy as an ontology-based method), evaluation is described in Table 5 (according Park et al. [25]).

Evaluating the results. It is related with the results. Evaluation by KE engineer and third-party experts is implicit in Stage 3 since ontologies built by other experts are considered. Domain experts evaluation is necessary in order to estimate how results of Stage 4 are accepted in the way Table 4 addreses it. Also, the limitation of evaluation against other KE experts to ontology comparison prevents *superhuman fallacy* (cf. Chandrasekaran's [10]) issues. Also, it could be interesting to use contingency tables in the evaluation against domain experts.

6 Conclusions

The idea of applying Ontological Engineering to SIS is supported by a number of security ontologies with different features and scope (see Blanco et al. [6] for a general vision of the field). The article shows how ontology creation from security reports—instead of selecting a standard security ontology—enables the use of formal methods assuring their safety, by clarifying processes and descriptions. As it was mentioned in the introduction, the goal is not to build (another) ontology on

security, nor to reproduce a standard method to extract one ontology from a document. The goal con-sists in exploiting the ontology creation process itself in order to clarify and revise security reports by reusing knowledge from a number of pre-existent security ontologies. Similar conclusions were obtained in other fields where similar semantic strategies were applied (Borrego et al. [7]). In fact, the strategy presented in the present work shares some tasks with NeOn methodology for the sce-nario where Reusing and Re-engineering non-ontological resources are necessary (Suárez-Figueroa et al. [35]).

The methodology presented in the article is not intended to be applied in documents explaining methodologies (e.g. Akrout et al. [1]). The methodology is related to the documentation of their pragmatic application at the operators level, with intra-organization scope. In fact, KE from incident reports authored by organization members would produce ontologies where the influence of human behaviour in SIS activities exists (e.g. in the population of flowchart instances) but it is not explicitly stated (as in others as, for example, Parkin et al. [26]).

The ontology is not the primary goal. The extraction of ontologies from SIS reports represents an excellent method to standardize and debug SIS knowledge; the application of ontology extraction (OE) methods is useful to reflect and refine SIS documentation and protocols. Security ontologies are built in the traditional fashion used in OE, while our approach is the reuse of OE methods to validate reports in IRD framework. Of course, an information security ontology should define the most important security issues, concepts and the relationships among them. Therefore, reports as those analysed here, should describe such elements. Thus, OE methods will produce ontologies, which are comparable with the former ones. The soundness of the report can be estimated, by means of this comparison, in order to induce refinements or reparations. Stage 3 is devoted to this task.

For example, it is interesting to compare the ontology with the Security Ontology (SO)[5] from Herzog et al. [16]. This ontology and the ontological elements, created during the strategy, com-plement each other with features as risk identification (from ours to the SO) and countermeasures analysis (from the SO to ours). It is evident that the countermeasure ontology from SO is richer than INTECO's [36], although it can be useful as addenda of the document.

In this article, the merging of the ontology with other—as for example that of Mace et al. [19]—is not considered. However, this would allow its expansion by analysing the relationship between human behavioral factors (and other concerns within information security management). The main reason for discarding it is that the proposed strategy is KE oriented, thus the inclusion of a relevant number of external concepts would distort the proper process of KE. A more suitable approach would be to assist the strategy with collaborative documentation tools from Knowledge Intensive Process (see Aranda-Corral et al.'s works[4, 5]). These tools bridge the knowledge gap between SWT experts and SIS experts.

If standardization efforts are part of the organization strategic plans, it could be interesting to consider in Stage 2 the relationship with security ontologies, in order to formalize (by means of a tool) the comparison with ISO 17799 standard for security (which involves ten security domains. See Microsoft's [20]). This particular relationship is important since it suggests the human operator, which tends to converge with international standards. In fact, ontology mapping plays a key role in this approach (as it is shown in Fenz et al [12]).

Lastly, it is worthy to mention that the strategy proposed in this article could be adapted to other Knowledge Intensive Methods for information spreading, as for example, in cross-enterprise collaboration approaches, particularly those where semantics can play a key role: business process management and knowledge externalization. Efforts for exploiting semantic web methodologies

[5] http://www.ida.liu.se/~iislab/projects/secont/

in the first approach represent a cornerstone to achieve true knowledge transmision and sharing [17]. In this way Business Process Modelling (BPM) in non OE scenarios could be enhanced with similar results than those from SIS field. In this case it is possible that the extracted ontology may be completed in order to be standard to BPM, enabling automated composition and planning of processes with robustness.

Acknowledgements

Partially supported TIN2013-41086-P project (Spanish Ministry of Economy and Competitiveness), co-financed with FEDER funds.

References

[1] R. Akrout, E. Alata, M. Kaâniche and V. Nicomette. An automated black box approach for web vulnerability identification and attack scenario generation. *Journal of the Brazilian Computer Society*, **20**, 1–16. Springer, 2014.

[2] J. A. Alonso-Jiménez, J. Borrego-Díaz, A. M. Chávez-González and F. J. Martín-Mateos. Foundational challenges in automated semantic web data and ontology cleaning. *IEEE Intelligent Systems*, **21**, 42–52, 2006.

[3] G. A. Aranda-Corral and J. Borrego-Díaz. Mereotopological analysis of formal concepts in security ontologies. In *Proceedings of the 3rd International Conference on Computational Intelligence in Security for Information Systems (CISIS'10)*, pp. 33–40, 2010.

[4] G. A. Aranda-Corral, J. Borrego-Díaz and Antonio Jiménez-Mavillard. Social ontology documentation for knowledge externalization. In *Proceedings of the 4th International Conference Metadata and Semantic Research*, pp. 137–148, 2010.

[5] G. A. Aranda-Corral, J. Borrego-Díaz, Juan Galán Páez and Antonio Jiménez-Mavillard. Emergent concepts on knowledge intensive processes. In *Proceedings of the 6th International Conference Computational Collective Intelligence (ICCCI 2014)*, pp. 282–291, 2014.

[6] C. Blanco, J. Lasheras, R. Valencia-García, E. Fernández-Medina, J. Ambrosio Toval Álvarez and M. Piattini. A systematic review and comparison of security ontologies. In *Proceedings of the 3rd International Conference on Availability, Reliability and Security, ARES 2008*, pp. 813–820, 2008.

[7] J. Borrego-Díaz, A. M. Chávez-González, M. A. Martín-Pérez and J. A. Zamora-Aguilera. Semantic geodemography and urban interoperability. In *Proceedings of the 6rd International Conference Metadata and Semantic Research*, pp. 1–12, 2012.

[8] J. Borrego-Díaz, A. M. Chávez-González, J. L. Pro-Martín and V. Matos-Arana. Specifying and verifying meta-security by means of semantic web methods. In *Proceedings of the International Joint Conference SOCO'14-CISIS'14-ICEUTE'14*, pp. 355–365, 2014.

[9] D. Catteddu and G. Hogben. Cloud Computing: benefits, risks and recommendations for information security. *Technical report*. European Network and Information Security Agency, 2009.

[10] B. Chandrasekaran. On evaluating artificial intelligence systems for medical diagnosis. *AI Magazine*, **4**, 34–37, 1983.

[11] S. Fenz and A. Ekelhart. Formalizing information security knowledge. In *Proceedings of the 2009 ACM Symp. on Information, Computer and Communications Security*, pp. 183–194. ACM, 2009.

[12] S. Fenz, G. Goluch, A. Ekelhart, B. Riedl and E. R. Weippl. Information security fortification by ontological mapping of the ISO/IEC 27001 standard. In *Proceedings of the 13th IEEE Pacific Rim Int. Symp. Dependable Computing (PRDC 2007)*, pp. 381–388, 2007.
[13] K. Geers. *Strategic Cyber Security*. CCD COE Publication, 2011.
[14] R. J. Glushko and T. McGrath. *Document Engineering - Analyzing and Designing Documents for Business Informatics and Web Services*. MIT Press, 2008.
[15] A. Herrero, M. Navarro, E. Corchado and V. Julián. RT-MOVICAB-IDS: addressing real-time intrusion detection. *Future Generation Computer Systems*, **29**, 250–261, 2013.
[16] A. Herzog, N. Shahmehri and C. Duma. An ontology of information security. *International Journal of Information Security and Privacy*, **1**, 1–23, 2007.
[17] Hanh H. Hoang, Jason J. Jung and Chi P. Tran. Ontology-based approaches for cross-enterprise collaboration: a literature review on semantic business process management. *Enterprise Information Systems*, **8**, 648–664. Taylor & Francis, 2014.
[18] W. Kim, O-R Jeong, C. Kim and J. So. The dark side of the internet: Attacks, costs and responses. *Information Syst.*, **36**, 675–705, 2011.
[19] J. C. Mace, S. Edward Parkin and A. P. A. van Moorsel. A collaborative ontology development tool for information security managers. In *Proceedings of the 4th ACM Symp. Comp. Human Interaction for Management of Information Technology, CHIMIT 2010*, p. 5, 2010.
[20] Microsoft Inc. *Enterprise Risk Management Models*. Microsoft, 2010.
[21] C. Minonne and G. Turner. Evaluating knowledge management performance. *Electronic Journal of Knowledge Management*, **7**, 535–662, 2010.
[22] National Academy of Sciences and Royal Society. *Cybersecurity Dilemmas: Technology, Policy, and Incentives: Summary of Discussions at the 2014 Raymond and Beverly Sackler U.S.-U.K. Scientific Forum*. National Academic Press, 2015.
[23] I. Nonaka and H. Takeuchi. *The knowledge-creating company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, 1995.
[24] D. L. Olson and D. Wu. *Information Security Management System for Microsoft's Cloud Infrastructure*. Springer Berlin Heidelberg, 2010.
[25] J. Park, W. Cho and S. Rho. Evaluating ontology extraction tools using a comprehensive evaluation framework. *Data & Knowledge Engineering*, **69**, 1043–1061, 2010.
[26] S. Edward Parkin, A. P. A. van Moorsel and R. Coles. An information security ontology incorporating human-behavioural implications. In *Proceedings of the 2nd International Conference on Security of Information and Networks, SIN 2009*, pp. 46–55, 2009.
[27] T. S. Mendes Pereira and H. M. Dinis Santos. An ontology based approach to information security. In *Proceedings of the 3rd International Conference Metadata and Semantic Research*, pp. 183–192, 2009.
[28] S. Sadvandi, N. Chapon and L. Piètre-Cambacédès. Safety and security interdependencies in complex systems and sos: Challenges and perspectives. In *Complex Systems Design & Management - Proceedings of the 2nd International Conference on Complex Systems Design & Management*, pp. 229–241, 2011.
[29] S. Sadvandi, N. Chapon and L. Piètre-Cambacédès. Safety and security interdependencies in complex systems and sos: Challenges and perspectives. In *Proceedings of the 2nd International Conference on Complex Systems Design & Management, CSDM*, pp. 229–241, 2011.
[30] A. Sarmah, S. M. Hazarika and S. Kumar Sinha. Security pattern lattice: A formal model to organize security patterns. In *19th International Workshop on Database and Expert Systems Applications (DEXA 2008)*, 1-5 September 2008, Turin, Italy, pp. 292–296, 2008.

[31] N. Shadbolt, K. O'Hara and L. Crow. The experimental evaluation of knowledge acquisition techniques and methods: history, problems and new directions. *International Journal of Human-Computer Studies*, **51**, 729–755, 1999.

[32] Shawn Riley. *Science of Cybersecurity Developing Scientific Foundations for the Operational Cybersecurity Ecosystem*. Centre for Strategic Cyberspace + Security Science, 2015.

[33] G. E. Smith, K. J. Watson, W. H. Baker and J. A. Pokorski II. A critical balance: Collaboration and security in the it-enabled supply chain. *International Journal Production Research*, **45**, 2595–2613, 2007.

[34] S. W. Smith and E. H. Spafford. Grand challenges in information security: Process and output. *IEEE Security & Privacy*, **2**, 69–71, 2004.

[35] M. C. Suárez-Figueroa, A. Gómez-Pérez and M. Fernández-López. The neon methodology for ontology engineering. In *Ontology Engineering in a Networked World*, pp. 9–34. Springer, Berlin, Heidelberg, 2012.

[36] J. Díaz Vico, D. Fírvida Pereira and M. A. Lozano Merino. *Identification and Reporting of Security Incidents for Strategic Operators. A Basic Guide for the Protection of Critical Infrastructures*. NICT, 2014.

[37] J. Díaz Vico, D. Fírvida Pereira and M. A. Lozano Merino. *The Operator Console. A Basic Guide to Critical Infrastructure Protection*. NICT, 2014.