KMapper – An Ontology-Based Knowledge-Mapping Application

Mrs. Lecocq Régine

2459 Boul. Pie XI Nord | 2459 Pie XI Blvd North
Québec, QC, Canada G3J 1X5
Regine.Lecocq@drdc-rddc.gc.ca
Regine.Lecocq@forces.gc.ca

ABSTRACT

This paper presents the advancements of a knowledge-mapping (k-mapping) ongoing research being conducted at Defence R&D Canada – Valcartier. In this research, the k-mapping concepts and techniques are investigated through the development and demonstration of a k-mapping alpha prototype application called "KMapper". The prototype has been developed and demonstrated for the benefit of the Canadian Forces and it is based on k-mapping underlying concepts as well as on a new multidimensional approach conceived at Defence R&D Canada. In this research k-mapping aims to enable the discovery, identification, localization, and access of essential pieces of knowledge called knowledge assets (KAs). These KAs, considered as critical to the commanders to ensure mission success, are obviously not all held within their own organisation. Instead KAs can be identified as originating from other military or non-military organisations having a stake in the complex situation being faced. The paper explains k-mapping concepts along with an innovative approach and a k-mapping alpha prototype application called "KMapper". Next, it discusses the promises of the approach and the application in response to some of the challenges faced by the Canadian Forces as a collaborative organisation. In relation to this, the paper introduces the concept of ontology and how ontologies are an intrinsic part of the KMapper prototype in order to carry its foreseen value. Subsequently, still with a special focus on ontologies, the paper discusses the rationales for it, the issues being faced and the solutions envisioned. Finally, some future work and research are proposed and conclusions are drawn.

1. INTRODUCTION

The Canadian Forces (CF) along with the coalition's military Forces "are now facing, to a greater extent than ever, challenging operating environments" [4]. Abroad or in Canada, the Armed Forces have to operate in more complex environments, which require "Forces that are combat-effective, but also highly mobile, adaptive, networked, sustainable and capable of operating in a Joint, Interagency, Multinational and Public (JIMP) context." [5] These intricate operating environments generate corollaries such as complexified technological solutions and overloads of very different pieces of knowledge/information coming from a large variety of sources. These complex environments also require from our Forces more sophisticated common operating pictures where these different pieces of knowledge/information – called knowledge assets (KAs) in this paper – are being held by numerous distinct groups and organisations. Therefore, Commanders at all levels have to compose with those different military and non-military organisations which hold portions of the answer to attain mission success. As the complexity of the situation and the number of involved organisations increases, it becomes very difficult for the militaries to understand the different domains related to a particular situation but also to identify which organisation holds which critical KAs. These situational KAs need first to be identified, located and made available to only then be exploited adequately by military personnel.

During the last few years, the Department of National Defence and the CF have started to investigate the field of knowledge-mapping (k-mapping) as one of the solutions to solve KAs discovery and localisation. The value of k-mapping concepts and techniques were investigated through the development of a k-mapping alpha prototype application named "KMapper". One of the focal objectives of the research is to demonstrate the value of a dynamic k-mapping capability in order to:

- Enhance sense-making of the situation by identifying the different KAs;

- Improve knowledge understanding by revealing the links between KAs;
- Support decision making process based on better understanding of the situation and the organisations involved; and
- Increase internal and external collaboration as knowledgeable individual and organisations are revealed.

This paper first presents, at a high level, the k-mapping alpha prototype application called "KMapper" along with its leading concepts, underlying multidimensional approach and technological features. Second, the paper introduces the concept of ontology and how ontologies are an intrinsic part of the KMapper prototype in order to carry its foreseen value. Finally, with a special focus on ontologies, the paper discusses the rationales for it, the issues being faced, the solutions envisioned and the future research and development related to the project.

As mentioned, the raison d'être of this research project resides in the growing number of situations where the CF have to collaborate with many different instances. While the k-mapping prototype application was demonstrated during a military exercise concerning national/public security situation within a Joint, Interagency, Multinational and Public (JIMP) context; the parallels are easy to draw with coalition operations.

2. KNOWLEDGE MAPPING FOUNDATIONS

From the literature, as well as from the field of practice, the specific term of "knowledge mapping" is commonly used to refer to completely different concepts originating, in some cases, from disconnected scientific fields. Despite such a dissent, k-mapping can be articulated along one of three main categories of approaches namely the conceptual approach, the procedural approach, and the social approach. As depicted in figure 1, the category named "conceptual approach" considers the value of differentiating and structuring concepts and terms. It usually relates to concept mapping [12], taxonomies, thesaurus or else ontologies and is of particular interest to ease learning or understand complex environments. From a high level standpoint, the category of procedural approach [7] envisions k-mapping as the mapping of processes where knowledge can either be the input, the output or else the sole objective of specific processes. The third category, named "social approach" [2] reflects on k-mapping from a social perspective; it considers the relationships existing between individuals as performed in social network analysis or link analysis, the localisation of key experts or else the flow of knowledge within organisations. As a consequence of such a discrepancy, when times come to undertake a k-mapping project, practitioners tend to respond to the requirements in the light of only one of those k-mapping approaches. Consequently the results of such projects tend to be disappointing as only limited values of k-mapping are considered or even worst the perspective/approach embraced is inappropriate to the pursued project objectives.

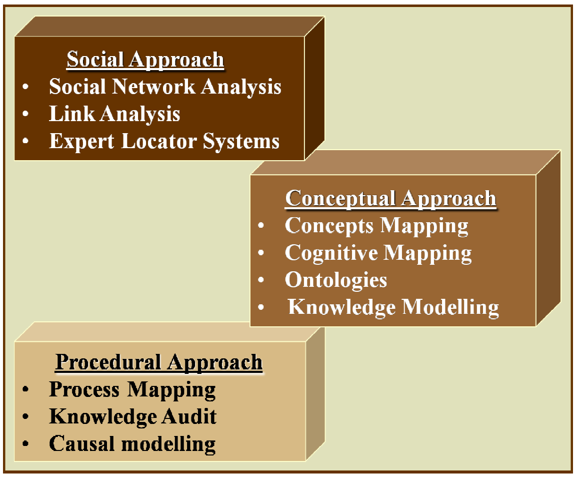

Figure 1: Three main k-mapping categories of approaches.

2.1. What is "knowledge-mapping" in this context?

In this research project, in order to respond to the CF context described previously, k-mapping is conceived as a process to discover, locate, and map specific KAs considered essential by individuals to perform their tasks. This research and its resulting k-mapping application called the KMapper, draw on a previous research project [8] converging towards a novel multidimensional k-mapping approach. The novelty of Defence R&D Canada's k-mapping approach resides in the fact that it combines the value of each one of the three categories of approaches previously described into a single one. Moreover, to this approach entitled the "Multidimensional K-mapping Balanced Approach" (MKBA); a fourth category named "Knowledge Artefacts" is added in order to support some information management requirements. Figure 2 represents the MKBA embedded in the k-mapping alpha prototype application. In it, the four different categories are presented as dimensions: the *Social Dimension*; the *Knowledge-Artefact Dimension*; the *Conceptual Dimension*; and the *Process Dimension*. Embracing such a multidimensional approach has several impacts on the work to be performed. Indeed, the multiple standpoints taken necessitate considering many domains of practice and expertise as for instance data, information and knowledge discovery; knowledge organisation as well as knowledge sharing as well as potentially the domains of ontologies and taxonomies and to some social aspects at the individual, group and organisational level.

2.1.1. Types of Knowledge Assets

This research project attempts to identify and locate two types of KAs (top portion of figure 2) held by distinct groups and organisations. The first type of KAs belongs to the *Social Dimension* and concerns tacit pieces of knowledge. This requires therefore identifying individuals, groups or organisations with specific expertise or experience. The second type of KAs belongs to the *Knowledge Artefacts Dimension* and relates to explicit knowledge such as incidents logs, documents, databases, websites, lessons learned, etc. From a technological stance, for these two dimensions, the k-mapping alpha prototype application uses the information gatherer services (explained in section 3.3.3) to automatically extract KAs and store information about them.

2.1.2. Organising and presenting the Knowledge Assets

Once the two first types of KAs discovered under the *Social* and the *Knowledge-Artefact* dimensions, relevant links are created and stored between the KAs and elements from the two other dimensions: the *Conceptual* and the *Process* dimensions (bottom portion of figure 2). Visually presenting the KAs along those two latest dimensions permits the users to comprehend the KAs within a meaningful context. It also provides new knowledge to the users by making apparent unexpected links between KAs and new concepts or specific process stages. Indeed, the *Conceptual Dimension* of the KMapper enables to present visually to the military user a domain-specific ontology with its related concepts and relations. By doing so, the end-users can immediately ascertain to which concepts a KA is related. Similarly, the *Process Dimension* covers the key processes being worked with by the targeted group of users. Whenever a specific stage of the process is being worked with, the KAs that should be prioritized in the context of that specific process stage are presented. Here again, the *Process Dimension* acts as a contextual element permitting the positioning of the KAs.

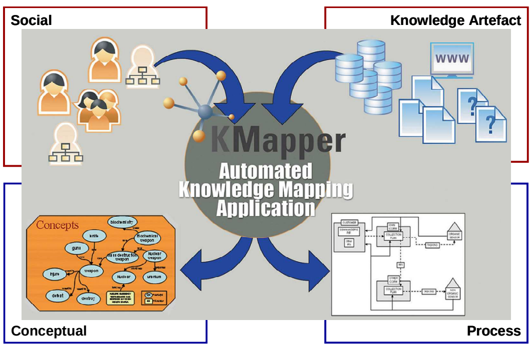

Figure 2: Multidimensional K-mapping Balanced Approach (MKBA).

3. THE KMAPPER – AN ONTOLOGY-BASED PROTOTYPE APPLICATION

Once this newly approach proposed, the research project aimed at demonstrating the value of k-mapping for the benefits of the CF through the instantiation of a k-mapping alpha prototype application named KMapper (figure 3). The KMapper prototype application mirrors the MKBA with its four key dimensions. Also, at the offset of the project, it was identified that usage of ontologies was of great value for the foreseen prototype.

For instance, the k-mapping alpha prototype application is dynamic and almost fully automated in its identification, localization and mapping of KAs. Such a characteristic is enabled by the information gatherer services which are based on a domain ontology acting as the application's backbone. The information gatherer services are further exposed in section 3.3.3. As another example, whereas trying to map the world would be a useless activity, the domain-specific ontology supports the search engines in order to retrieve information about KAs relevant to the end-user in the context of the situation being faced only. This situation-focused activity provides military end-users – accomplishing different tasks – with a shared context for common actions. Furthermore, the ontology structure and content also feed the "Knowledge Inference" service. This service with its set of rules is essential to discover and gain new relevant pieces of knowledge based on the ones already accumulated. These examples are explained in more details in the next sections and some of them particularly apply to coalition context.

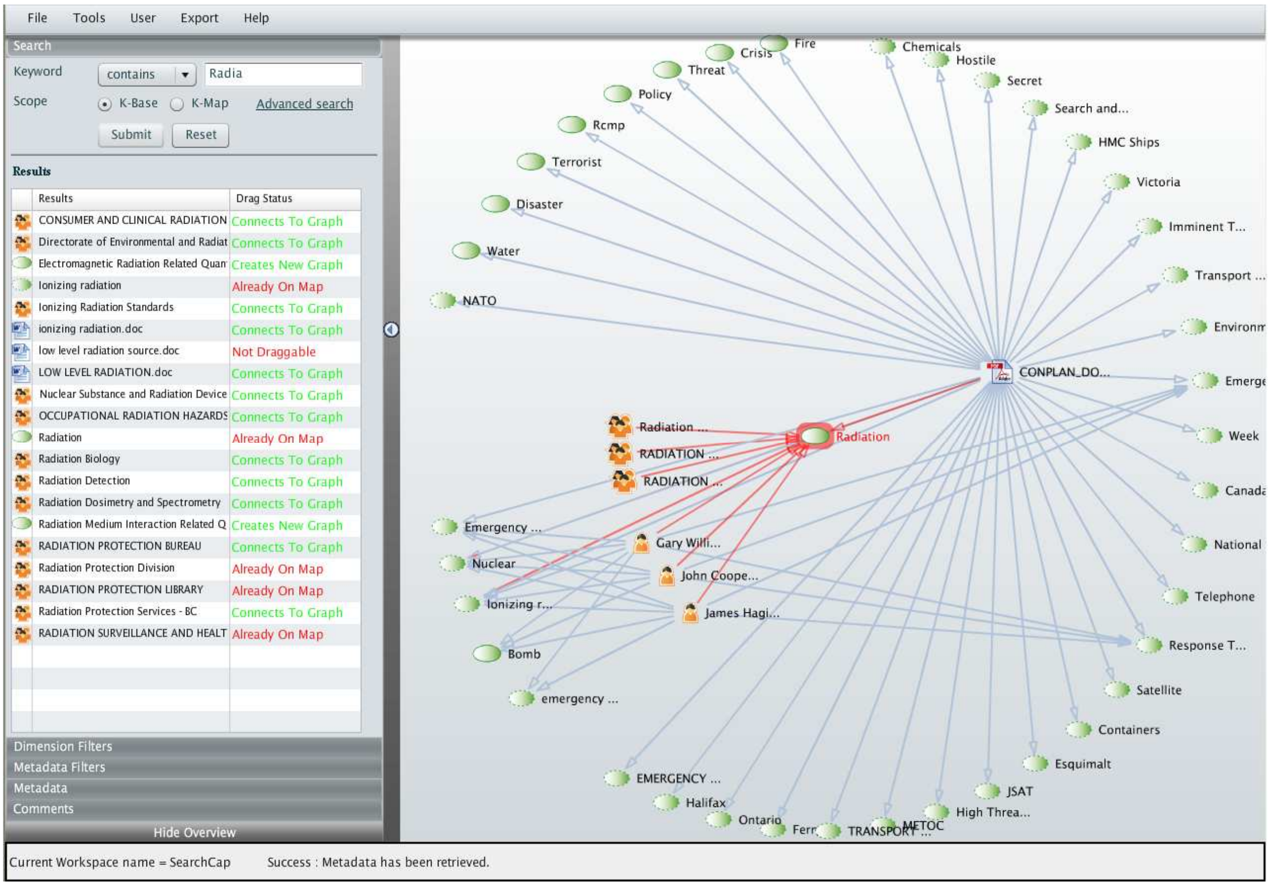

Figure 3: KMapper screen shot.

3.1. What are ontologies?

As exposed by Gómez-Pérez et al. [6] many definitions of what is an ontology exist. Some of them are more focused on the earliest philosophical aspect of it while others consider the newly technological characteristics specified under the auspices of the knowledge engineering community. One of these definitions pertains particularly to the context of this research and is given by Uschold and Jasper [13]:

> *"An ontology may take a variety of forms, but it will necessarily include a vocabulary of terms and some specification of their meaning. This includes definitions and an indication of how concepts are inter-related which collectively impose a structure on the domain and constrain the possible interpretations of terms."*

Due to the technological aspects of the KMapper the definition from Schreiber [11] would also apply:

> "An ontology *provides the means for describing explicitly the conceptualization behind the knowledge represented in a knowledge base."*

3.2. How are ontologies instantiated in the KMapper?

The KMapper is composed of two types of ontologies. Figure 4 depicts the first ontology that constitutes the prototype, the "KMapper core ontology", and supports the application itself. This ontology permits the definition and description of concepts and their relationships related to the KMapper application, as well as the application's own structure, dimensions, etc. It is a permanent part of the application and is required by the system to organise the KAs under the appropriate dimensions and with adequate properties. The Protégé ontology editor was used in the prototype in order to manage the ontological aspects of the application.

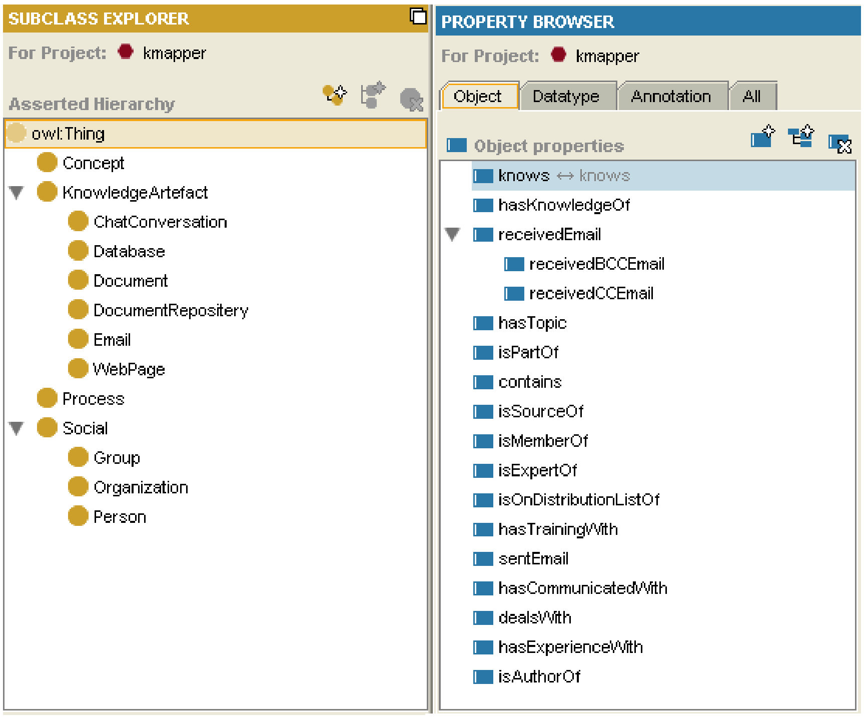

Figure 4: KMapper core ontology.

In addition to the core ontology, the system also requires a domain ontology (lower portion of figure 5). The domain ontology relates to the *Conceptual Dimension* and can potentially be replaced or augmented with another ontology of interest. The domain ontology is considered as the backbone of the KMapper as it supports the search engines in retrieving information about KAs of significance to the end-user. First, using the information gather services, the application searches data sources to extract KAs, subsequently it is also around the concepts of the domain ontology that the extracted KAs are organised and linked to one another. The current version of the KMapper is supported by a domain ontology largely based on a chemical, biological, radiological, nuclear, and explosive (CBRNE): the AKTiveSA[1] ontology.

[1] The AKTiveSA Ontology was built by the School of Electronics and Computer Science, University of Southampton, UK.

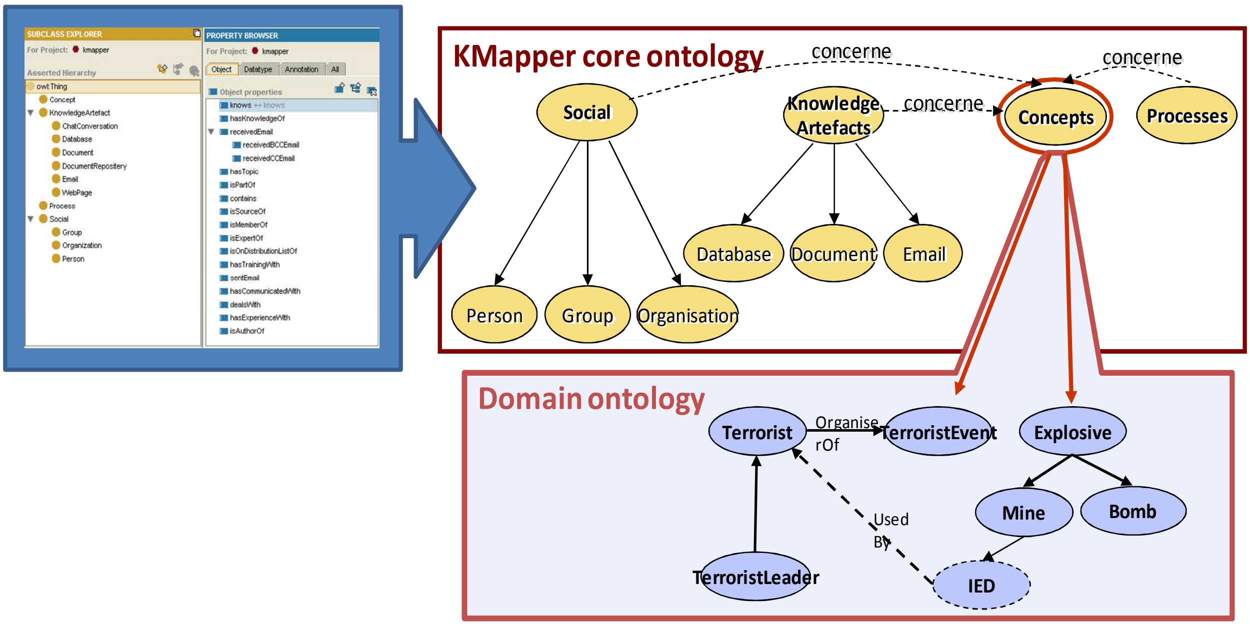

Figure 5: KMapper ontologies.

3.3. Why an ontology-based application?

During the last two to three decades, the domain of ontologies has been more accurately investigated from a computer science perspective under the auspices of the knowledge engineering community. This paper will not attempt to dissect the domain of ontologies but rather share the rationales to, as well as lessons learned from, generating such an ontology-based application.

3.3.1. Reason 1: To have a prototype corresponding to the selected approach

As described previously, the MKBA is based on four dimensions, amongst which one is the *Conceptual Dimension*. The *Conceptual Dimension* is particularly well represented with ontologies. The *Conceptual Dimension* draws from the domain of concept maps and in the KMapper, the concept dimension is based on ontologies. Concept maps have commonalities with ontologies; they equally relate to concepts and the relationships between them and in both cases, these relationships go beyond the parent/child relationships of taxonomies. On the other hand, initially, in the concept maps, concepts were organised in a hierarchical way [10], which is not a requirement for ontologies. Concept maps are of great usage in the educational and learning domain [10]; it eases the understanding of its reader and helps in retaining more easily the information as it corresponds to some stages of memory building. In the same vein, this *"Conceptual Dimension"* aims partly at enabling the understanding of the situation by the end-users of the KMapper.

Ontologies also provide a possibility towards machine readable representation of knowledge, which is very important as we built a prototype application. For the purpose of building the prototype, amongst the different available formats, OWL-DL was retained. OWL is one of the most frequently used web ontology language and therefore hopefully creates some stability for the prototype in the future. Also, generating an ontology is a very cumbersome task and it is usually of good practice to start from existing ontologies, which are in many cases expressed in OWL. Finally, as the research project envisions integrating some reasoning capability to the prototype, the OWL-DL variant of OWL was selected for that motive.

A few challenges emerged due to this technological choice. First, as the concepts and their relationships are expected to be visually presented to the end-users, the prototype requires some visualisation capabilities but more importantly, the visualized "name tags" necessitates to be expressed in natural language in order to be understood by the end-users. To do so, either the ontology needs to be expressed directly in natural language or else, the system necessitates corresponding tables. In order to be machine readable, ontologies normally do not require to be expressed in natural language and such constrains are somehow opposite to the essence of ontologies. This issue is partly solved by producing a domain ontology based on natural language but this solution imposes some intensive verifications of the ontology and also precludes from using other existing ontologies of potential interest but expressed in a different format. Other alternatives will need to be considered in the future.

3.3.2. Reason 2: The prototype should be context driven

One of the main outputs of the k-mapping application is to provide end-users with a contextualized visualisation of the discovered KAs as well as the links existing between them. Without trying to become an expert of each type of issue being faced, the mapping of KAs can at least start building the first blocks which will enable the treatment of specific upcoming information within short delays. Whereas mapping the world would be an impossible and useless task, using ontologies enables us to restrain the domain that requires to be mapped and therefore allows a more focused search on potential events in specific situations. For instance if there is an indication of the potential presence of key terrorist groups during an event, the individual might want to start consulting information about those groups as for instance their preferred types of attack and what it requires. This avoids having to map the world KAs about all topics. Another benefit from the ontology resides in the fact that by identifying the context surrounding the specific identified KA, additional information are brought to the end-users. As depicted in figure 6, being able to present the KAs in the context of the concepts and how they relate to one another, can bring knowledge to the end-user first by identifying the existence of specific concepts then by understanding how they relate to one another.

Figure 6: Visualising KAs in relation to concepts.

3.3.3. Reason 3: The prototype should be as automated as possible

As described by Ackerman [1] systems like experts locators usually lose their impact rapidly when having to update them manually; a certain level of automation is required. In the KMapper in order to be automated, one essential component of the KMapper is the information gatherer services, which rely entirely on the domain ontology. As depicted in figure 7, the information gatherers use directly the concepts and instances from the ontology to parse different data sources in order to automatically locate potential KAs. For instance, an information gatherer can identify one or more concepts from the ontology in a web page. In that situation the web page is tagged with metadata labels, which are stored in the KMapper knowledge base to be subsequently presented to the end-user in relation to its KA.

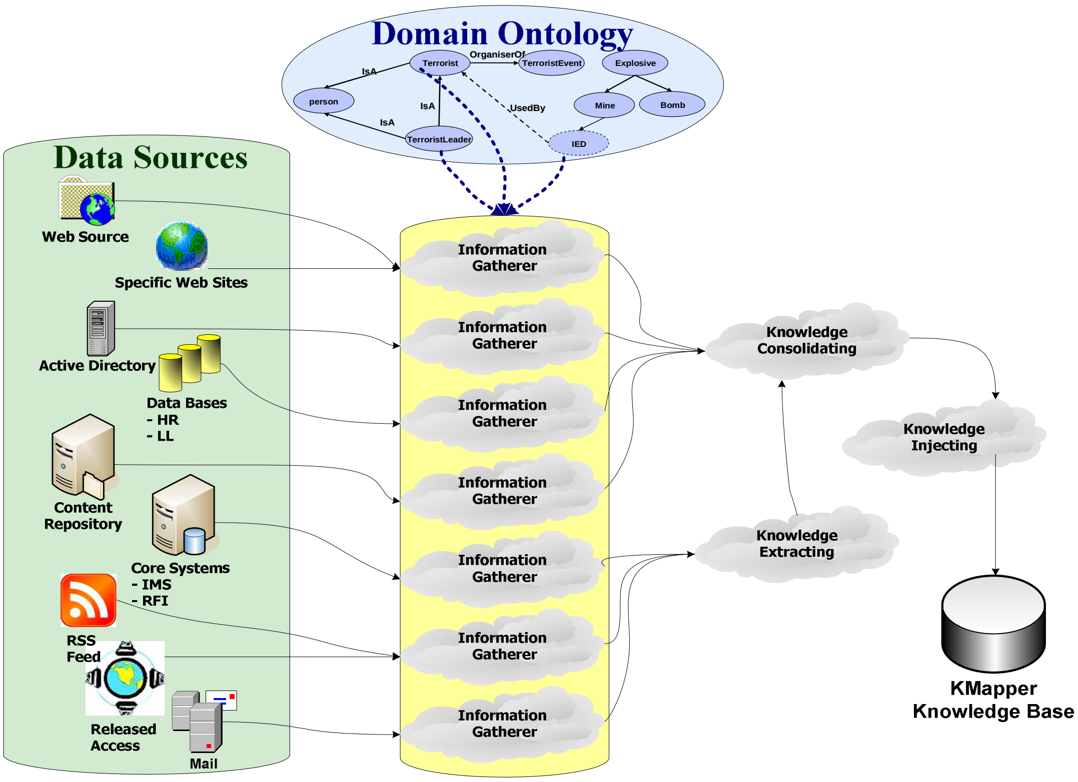

Figure 7: Information gatherer services using domain ontology.

The information gatherer services works quite well in the KMapper and it is actually one of its main values. Two challenges were however faced and solved. The first one resided in the need to parse using not only the concepts from the ontology but also the instances, including instances added manually by the end-users. The second issue relates to the one already exposed about the requirement to have an ontology using natural language. Indeed, as the information gatherers were parsing data sources including unstructured ones like texts, the concepts from the ontology required then to be expressed using natural language in order to effectively find KAs.

3.3.4. Reason 4: The prototype should be dynamic

This type of k-mapping application aims at pointing up-to-date information and knowledge to the end-users. If the prototype is not built in a dynamic way and constantly refreshed, as soon as the end-user will realize that the information is obsolete, he/she will lose confidence and stop using the application. This challenge is intimately related to the information gatherer services as the automated identification of the KAs also permits to develop the dynamic aspect of the KMapper. In fact, on a recurrent basis a scheduler service enables an automated and incremental update of the KMapper knowledge base and therefore of the information provided to the end-user. The scheduler is set-up differently depending on the data source being parsed using the ontology concepts. The data sources do not all require to be parsed at the same time but rather accordingly to their refreshment pace. The KMapper, through an administrative control panel, also provides the possibility to refresh the information upon demand.

This aspect of the application works quite well but from an ontological perspective one of the challenges, already exposed, is to enable the parsing with instances. Another challenge is to facilitate just-in-time insertion of new concepts and instances by the militaries. This is actually also facilitated through the administrative panel. These newly inserted concepts are then taken into consideration for future updates of the knowledge base. Consequently, new requirements emerge in order to ensure that the main ontology will not be contaminated with these new instances and concepts. Indeed, an ontology is a consensus and when it is time to update it, it is important to follow a clear and strict process.

3.3.5. Reason 5: The knowledge should be easily shared among partners

The KMapper prototype is built in order to find KAs residing within the CF but also outside the borders of the organisation; such an objective definitely beneficiates from the usage of ontologies. In fact, ontologies are supposed to enhance interoperability[2] and when having to parse data sources from other organisations, using well defined ontologies should allow overcoming the terminology differences between organisations while focusing on the exact meaning of the concept itself.

3.3.6. Reason 6 – The prototype should be generic

From a technological standpoint now, the usage of domain ontologies increases the generic aspect of the k-mapping application. The different domain ontologies can be placed or replaced in the KMapper, just like cartridges and therefore permit to use the KMapper in very different context. The current KMapper prototype only permit the management of one domain ontology but the situation being faced by the militaries are usually manifolds and a k-mapping application should also beneficiate from the simultaneous combination of different ontologies equally related to the situation.

4. NEXT STEP, FURTHER RESEARCH AND DEVELOPMENT

The current KMapper prototype, which has been developed in the light of operational environments, is being integrated and revisited to respond to intelligence domain activities. More specifically, it is being integrated into a "Multi-Intelligent Tool Suite" (MITS) based on an Intelligence S&T Integrated Platform (ISTIP) responding to Service Oriented Architecture (SOA) principles. To perform this, it requires revising some of the underlying concepts of operation as well as the application functions as for instance, the automated reporting service, which was built in the system in order to respond to military operational requirements.

From a technological viewpoint several enhancements are planned. At first there is a need to improve the inference and filtering service. These are key elements in order to focus only on KAs pertinent to end-users and therefore helping the reduction of information overload. Secondly, there is a requirement to identify more sophisticated visualisation algorithms, specifically algorithms that could enable a more appropriate visualisation of the ontology per se. Thirdly comes the necessity to integrate time and spatial components, this will be of specific interest for the military community. Also, the end-user should be more informed about the reliability of the KAs identified. The KMapper alpha prototype was built with three of the four dimensions; the *Process Dimension* was investigated in a parallel research in the maritime domain. Therefore, the *Process Dimension* will need to be coupled to the remaining part of the system. In the domain of ontologies, some improvements are also required; in the future, the KMapper should be able to use different format of ontology languages for instance. Finally, current research in social network analysis and patterns of life recognition are also considering leveraging some concepts from k-mapping domain as well as, under the ISTIP, services from the k-mapping alpha prototype application.

[2] While ontologies were envisioned as such, some authors have doubts about the attainment of such a capability. These doubts are specifically based on the miss layering issue between the different format of ontologies [3].

5. CONCLUSION

The CF, like the allies, have to respond to very complex situations where mission success relies on their ability to take into consideration and collaborate with other instances. These instances can be other governmental organisations, allies or local representatives, each of which can hold critical KAs with respect to the situation being faced. Being able to identify and locate those KAs is not a trivial task but it is critical in coalition operations. Based on the results of this ongoing research project, leveraging the value of k-mapping concepts for this purpose seems to hold some promises as do the usage of ontologies in the envisioned technological solution. During the development of the prototype application, using ontologies imposed many constrains but in some specific context like the one being experienced here, it remains a valuable solution. Finally, in the computer science field, research in the domain of ontologies is abundant but much remains to be performed before being able to lever easily their full potential.

6. REFERENCES

[1] Ackerman, M. et al. (2003), *Who's There? The Knowledge-Mapping Approximation Project*, In MIT Press, Sharing Expertise: Beyond Knowledge Management, pp. 159-178, Cambridge

[2] Cross, R. et al. (2002), A bird's eye view: Using social network analysis to improve knowledge creation and sharing, IBM Institute for K-based Organizations

[3] De Bruijn, J., Polleres, A., Lara, R., Fensel, D. (2005), *OWL DL vs. OWL Flight: Conceptual modeling and reasoning for the Semantic Web*, In Proceedings of the 14th International World Wide Web Conference, WWW2005, Japan

[4] Defence R&D Canada. (2006), Defence S&T Strategy - Science and Technology for a Secure Canada

[5] Department of National Defence and Canadian Forces (DND/CF). (2007), *Land Operations 2021 - Adaptive Dispersed Operations*

[6] Gómez-Pérez, A., Fernández-López, M., and Corcho, O. (2004), *Ontological Engineering*, Springer-Verlag, London

[7] Kang, I. et al. (2003), *A framework for designing a workflow-based knowledge map*, Business Process Management Journal; Volume 9 No. 3

[8] Lecocq, R. (2006), *Knowledge Mapping: A Conceptual Model*, DRDC-RDDC Valcartier TR2006-118

[9] Lecocq, R. (in progress), KMapper: a knowledge mapping alpha prototype application, DRDC Valcartier

[10] Novak, J. D. and Cañas, A. J. (2006/2008) *The Theory Underlying Concept Maps and How to Construct and Use Them*, Technical Report IHMC CmapTools, Florida Institute for Human and Machine Cognition: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf

[11] Schreiber, ATh., Wielinga, B.J., and Jansweijer, W. (1995), *The KAKTUS view on the 'O' World*, In Skuce, D. (ed) IJCAI95, Workshop on Basic Ontological Issues in Knowledge Engineering, pp. 15.1-15.10

[12] Trochim, W.M.K. (2002), *Concept Mapping*, http://www.socialresearchmethods.net/kb/conmap.htm

[13] Uschold, M. and Jasper, R. (1999), *A Framework for Understanding and Classifying Ontology Applications*, IJCAI'99 Workshop on Ontology and Problem Solving Methods: Lessons Learned and Future Trends. Stockholm, Sweden