MindCollect: An Information Needs Management System for Coalition Situational Awareness, Sensemaking, Decision Making, and Mission Integration

David Kamien (Mind-Alliance Systems), **Michael Kelly** (Mind-Alliance Systems),
Tope Omitola (Electronic and Computer Science, University of Southampton),
Gary Wills (Electronic and Computer Science, University of Southampton)
david@mind-alliance.com, mike@mind-alliance.com, t.omitola@ecs.soton.ac.uk, gbw@ecs.soton.ac.uk

ABSTRACT

An efficient mechanism for bottom-up elicitation from domain experts is a key enabler for producing meaningful information requirements specifications. This paper describes "MindCollect," an information-requirements management system that consists of: (a) An Information Needs Profiling Module for Q&A-based elicitation and gathering of operational user tasks and associated information needs; (b) Answer dataset management; and, (c) Automated formulation of coherent task narratives (aka user stories) using techniques inspired by Natural Language Generation approaches. MindCollect has proven efficient and effective in the bottom-up elicitation of requirements from domain experts. The use of MindCollect for determination of requirements in a civil emergency management project is described. Finally, we describe future research directions, including the potential for leveraging Linked Open Data.

Author Keywords:

Commander's Critical Information Requirements, Coordination Information Requirements, Common Operating Picture, Essential Elements of Information, Human-Computer Interaction, Information Exchange Requirements, Information Products, Information Requirements, Information Requirements Management, Information Sharing, Intelligence Surveillance & Reconnaissance, Knowledge Elicitation, Linked Open Data, Mission Integration, Mission Threads, Natural Language Generation, Planning Information Needs, Priority Intelligence Requirements, Semantic Web, Situational Awareness Requirements Management, Q&A, User Experience Design.

INTRODUCTION

Intelligence communities, military coalitions and networks of organizations that engage in high-stakes crisis and disaster-response endeavors need a profound understanding--and sometimes a common/shared understanding--of the data, information, intelligence, expertise, and data analytics that people and organizations need to access, report, share, and incorporate into their analysis, planning, decision-making, collaboration and coordination. This understanding is vital for enhancing doctrine, procedures, training, the design of systems (e.g., for decision support, planning, intelligence, and situational awareness), operational planning, and processes such as intelligence production. Determining in advance which sources and specific informational content a role needs in order to produce a document, make a decision, generate a forecast, perform a task, or solve a problem can accelerate data gathering and free up time better spent otherwise (e.g., on value-add analytic thinking), which can enhance decision quality and consequently, mission execution.

Establishing where any given person or group should turn for the most relevant, accurate, and valuable inputs and knowledge support is immensely challenging given the complex variety of missions, the proliferation of procedural doctrine and information resources, and the diversity of systems, people and organizations in a coalition network.

The Need For Information Requirements Management Software

Information requirements management software tools can help organizations efficiently gather the data needed to specify information requirements in an online format so they can be analyzed to identify issues and suboptimal information flows that undermine decision making, coordination, and situational awareness. The data can also be used to determine if procedures embody best practices; and as inputs to other artifacts such as visualizations and checklists, process maps and operational views that provide an integrated understanding of how missions and processes fit together across partners [2].

Profiling information needs is also important because it yields valuable role-specific institutional knowledge. Additionally, the data can be used to create job previews and guides that specify the information flows and learning resources associated with roles so organizations can hire smarter and make new staff productive more quickly.

Information requirements can be used to ensure that people and mission partners have the capabilities needed to access essential information inputs. If-then interdependencies and information flows between tasks, roles, and organizations (e.g., notification procedures needed to de-conflict operations) can be examined, along with inconsistencies between plans and procedures. To focus improvement efforts, an organization might examine issues of each type across all scenarios, or for a particular scenario, process, or organization, including unconfirmed information inputs between tasks, roles, and organizations. This analysis can yield insights that enable organizations to be more strategic, deliberate, and efficient about putting policies, procedures, agreements, systems, training, and services in place to satisfy unmet information needs and close capability gaps, and enhance cross-agency information flow.

MindCollect is an information requirements management system that has been designed to be usable by every individual or group regardless of modeling skills. These include domain experts—commanders, operational planners, and non-military partners, etc.—who understand the information requirements associated with their roles, context, and specific situations in an operating environment

PAST WORK

Military and crisis management professionals use a wide variety of concepts that relate to information requirements, including: Situational Awareness (SA), Common Operating Picture (COP), Information Exchange Requirements (IER), Information Requirements Management (IRM), Priority Intelligence Requirements (PIR), Intelligence Surveillance & Reconnaissance (ISR) Collection, Coordination Information Requirements, Essential Elements of Information (EEI), Commander's Critical Information Requirements (CCIRs), and Mission Threads.

Information requirements per se have been studied extensively in various fields, including Management Information Systems (MIS), national and open-source intelligence, and Command and Control-related enterprise architecture. To our knowledge, neither Enterprise Architecture nor emergency response planning support tools have captured information requirements directly from domain experts in the field, let alone done so in a structured, machine-readable format that could be programmatically leveraged by matching them to information resources [3].

For example, in the NATO enterprise architecture framework, the NOV-3 artifact, "Operational Information Requirements" subview addresses the need to identify and describe all information exchanges that make up all information needlines between operational nodes, i.e it identifies who exchanges what information, with whom, why the information is necessary, and with what quality the information exchange must occur." [4]

In the crisis management domain, Synch Matrix, a PC-based decision and planning support tool developed by Argonne National Lab, provided a system for optimizing the planning, exercising, and implementation of emergency response plans, taking into account the interaction and activity flow of a plurality of independent organizations whose actions and decisions will affect the actions and decisions of other organizations in the emergency response process. The focus of this tool, however, was not on information requirements including the assurance that they are matched to resources.

Mind-Alliance Systems introduced key requirements for a software system for information sharing planning and collaboration modeling [5,6]. A proprietary software tool called Channels mapped information flows between tasks in planning scenarios and produced information-sharing procedures that specified information inputs and outputs for each assigned task and role involved in the plan [7]. Information sharing planners from different organizations determined what information needed to be shared; who would communicate it and by what means of transmission; how quickly information needed to be received; and classification levels, and governing policies based on inputs from interviews with domain experts. Visualizations in Channels depicted the consequences of failing to share information on dependent tasks and goals, total elapsed time for an element of information to propagate through a chain of tasks, and tasks that were incompletely specified [8]. Figure 1 shows the developer interface in Channels.

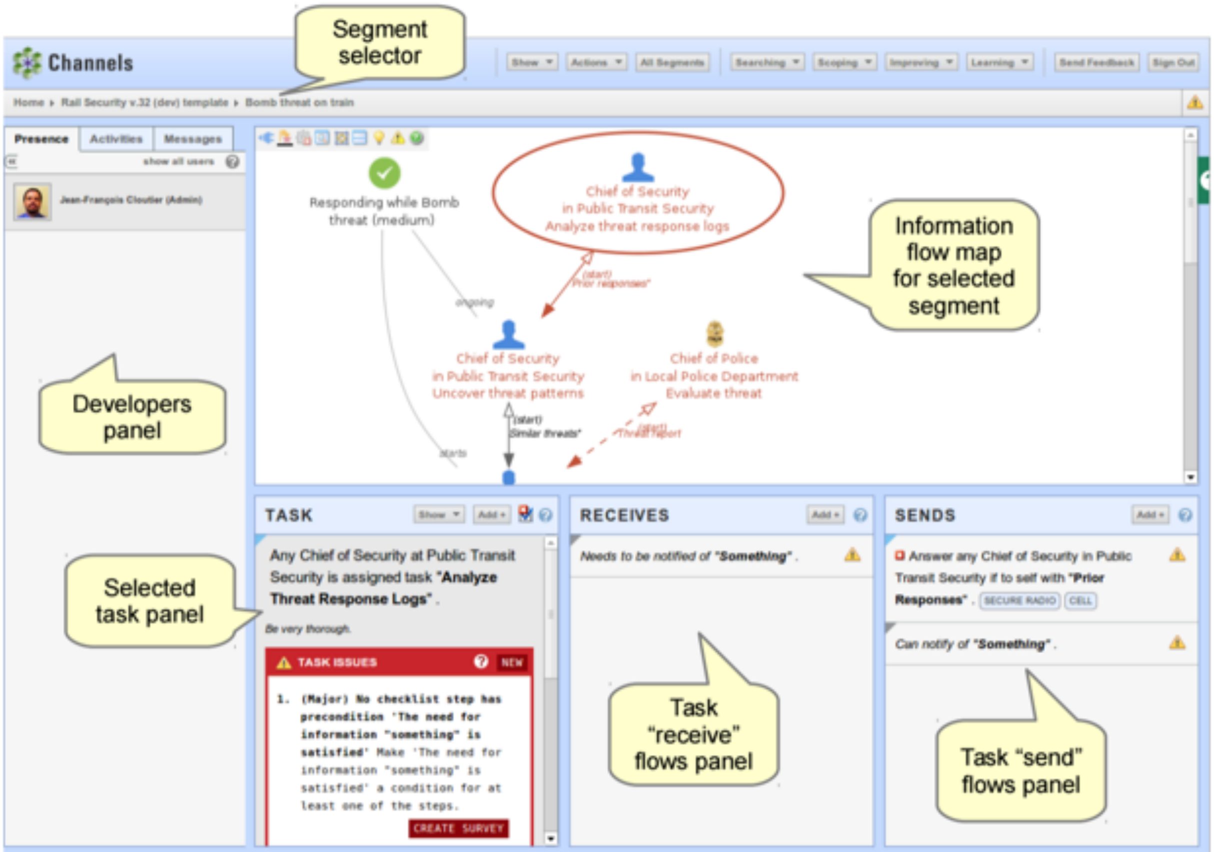

Figure 1: Screenshot of developer interface in Channels

Building collaboration models with Channels generated valuable insights into processes and detected potential issues that could derail planned collaboration. For example, in a 2010 information sharing planning project focused on preventing a terror attack on a major East Coast city, it became evident, through the use of Channels, that participating organizations intended to report suspicious activity incidents to multiple federal and regional agencies, and not to each other, and how specific delays in "connecting the dots" would result in blind spots and vulnerabilities. Incorporating Channels into a 2011 project for a state emergency management agency illustrated how a particular role in the Emergency Operations Center—the Watch Officer—would be overloaded with planned information and information requests and that the tasks assigned to that role had to be partially reallocated. In a 2011 Advanced Research Workshop, delegates from 10 NATO members and partner countries agreed on the need to systematically plan information sharing in a security context [9].

Although these and other projects highlighted the value of a systematic approach to analyzing information requirements and exchanges, it became evident that engaging Channels-trained consultants to interview people to elicit information about their information needs or training staff to use the software imposed excessive costs. Agencies involved in homeland security and emergency management would not invest the significant amount of time required to learn to use the complex Channels software system, as well as conduct interviews with responders and domain experts to acquire the needed data. With easy-to-use mobile apps becoming commonplace, users had little time or patience for software that required significant amounts of training and the learning of new frameworks or methodologies.

Another fundamental issue was that Channels was designed for centralized or top-down joint planning by a group of organizations needing to agree on the focus and scope of a planning scenario and the associated "Information Sharing Plan" or "Collaboration Plan." In practice, the scope of the information-sharing requirements could not be "nailed down" and constrained enough for a plan to ever be considered complete. Even a partial modification to the underlying planning scenario would require revisiting and updating the set of existing information requirements -- which meant consultants/analysts returning to the domain experts for further interviews and the clarification and validation of the information exchange requirements.

After years of Channels development, the need for a simpler system to profile information needs became increasingly clear. Unlike tools for enterprise architects, developers, business analysts, and professional emergency operations planners, a next-generation information requirements and exchange management system had to be usable by every individual or group regardless of modeling acumen. These include domain experts--commanders, operational planners, and non-military partners, staff at operations centers and warfighters in the field--who understand the information requirements associated with their roles, context, and specific situations in a mission and operating environment (COIN, HADR, Stability Operations). These findings were confirmed after an experimental project focused on using Channels to model Deployable Communication Information Systems (DCIS)-related operations at NATO Allied Command Transformation.

With the realization that User-Centered Design, Human Computer Interaction (HCI), and User Experience Design (UX) are fundamental requirements for information-requirements management software Mind-Alliance rebooted the development effort and began creating the MindCollect system with a Q&A-based dialog-like interface (described below) to lower the usability barrier.

SYSTEM DESCRIPTION

MindCollect consists of: (a) An Information Needs Profiling Module for Q&A-based elicitation and gathering of operational user tasks and associated information needs; (b) Answer dataset management; and, (c) Automated formulation of coherent task narratives (aka user stories) using techniques inspired by Natural Language Generation approaches. These narratives are user-friendly artifacts, which provide essential context for understanding the expressed information needs.

The MindCollect web application enables users without formal modeling skills to profile information needs and requirements related to situational awareness, sensemaking, decision making, and mission integration. To profile information needs and their context, users engage in a simple Q&A process. The MindCollect metamodel semantically interrelates the conceptual elements in the questions and answer datasets.

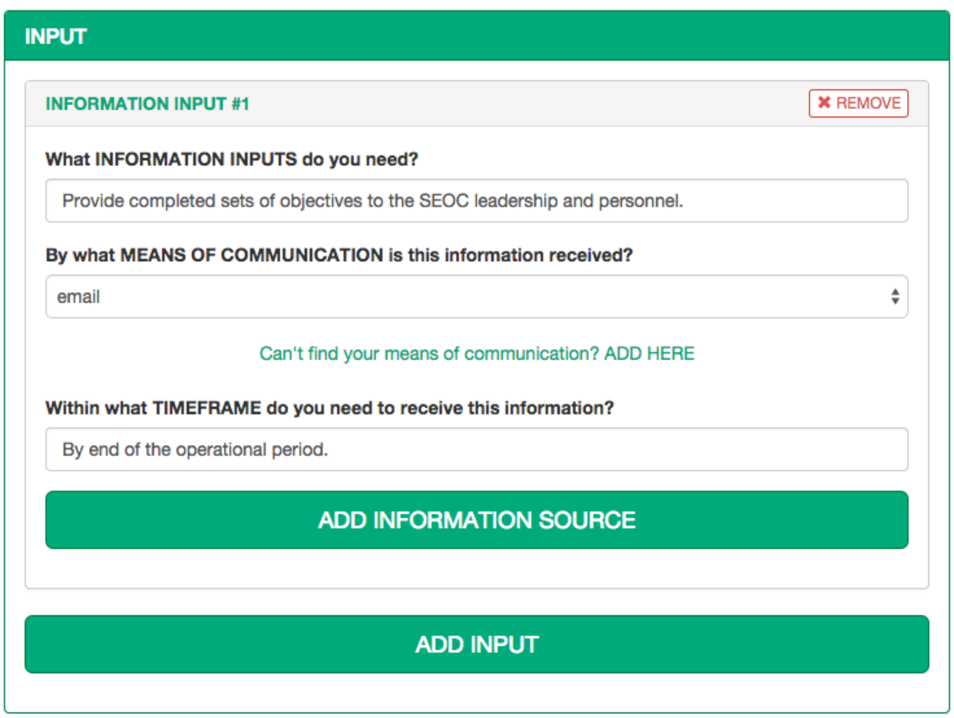

Figure 2: Input Q&A interface in MindCollect.

One of the goals of MindCollect is to provide an intuitive medium where non-technologist domain experts will be able to specify requirements by answering questions and then have the system formulate natural language narratives in English prose to facilitate validation and user engagement. Questions for profiling tasks and associated information requirements included:

- Do you perform this task during Phase X, Phase Y, or Phase Z of the PROCESS?
- Which ROLE in the above organization performs this task?
- What INFORMATION INPUTS are needed to execute this task?
- What is the SOURCE ORGANIZATION for this information?
- What is the SOURCE ROLE for this information?
- By what MEANS OF COMMUNICATION is this information received?
- What ISSUES impact your ability to access or use this information?

To be useful, information requirements need to be associated with a context. For example, a user can associate an information requirement with a specific task that is part of a process. When asking someone else to profile their information needs, a user can frame the request by specifying context factors (e.g., “What information do you need when you perform task X in response to event Y?”). To organize and enable querying and faceted sorting of information requirements, users have the option of tagging and associating them with contextual factors, such as business processes, decisions, tasks, goals, events, roles, projects, plans, documents, IT systems, and forecasts.

Natural Language Generation

Natural Language Generation (NLG) is the natural language processing task of generating natural language from a machine representation system such as a knowledge base or a logical form such as a template. Formally defined, an NLG system accepts a **<S,G,U,N>** tuple, where **S** is the knowledge source, **G** is the communicative goal, **U** is the user model, and **N** is the discourse model or narrative style. This means that “You know S and you want to say G to U, using the style of N” [10].

Simple examples include systems that generate form letters. These do not typically involve grammar rules, but may generate a letter to a consumer, e.g., stating that a credit card spending limit was reached. More complex

NLG systems dynamically create texts to meet a communicative goal. The typical stages of natural language generation, as proposed by Dale and Reiter [11] are:

- Content determination: Deciding what information to mention in the text.
- Document structuring: Overall organization of the information to convey. For example, deciding to describe the areas with high pollen levels first, instead of the areas with low pollen levels.
- Aggregation: Merging of similar sentences to improve readability and naturalness. This may not be necessary if sentences are short.
- Lexical choice: Putting words to the concepts.
- Realisation: Creating the actual text, which should be correct according to the rules of syntax, morphology, and orthography.

In NLG, the realization of the generated sentences (or narratives) depends on content determination and document structuring for the communicative goal to be successful.

The free-form inputs from domain experts in Vermont were serialized into JSON format, giving it a semi-structured format. The Dialog Manager of MindCollect was employed to turn users’ inputs into English prose, providing users with the ability to understand, coordinate, and communicate with other team members. Figure 3 shows a natural language paradigm synthesized by the software.

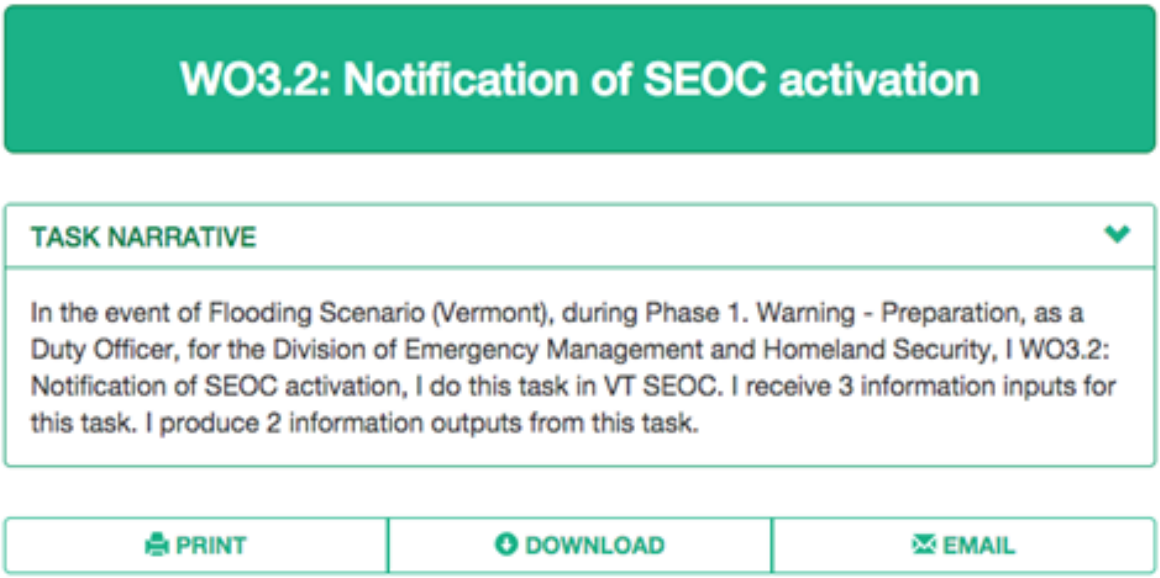

Figure 3: Task Narrative in MindCollect software.

Information Requirements Issues Management

MindCollect displays the information requirements that are unmet (information gaps) and have issues so a business process improver can plan improvements and track the effect of solutions that get deployed.

Information requirements profiles can include a question that enables users to indicate issues, such as the fact that they do not know the source of needed information. Additional questions can be used to indicate whether this issue adversely impacts task performance, productivity, and decision quality. Issue categories include: Accessibility/ Sourcing, Timeliness, Search, Traceability, Maintainability, Speed, Comprehensiveness, Accuracy, Conciseness, Correctness, and Currency.

Once an organization has identified information requirements that are not being met, information exchange requirements that have not been confirmed by counterparties, and other types of issues, they can sort them using the context factors, prioritize, and then plan solution interventions to address them. Personnel working on improving business processes, (e.g., KM professionals, IT, and business managers) can review information requirements with issues and then develop solutions to address them. Leaders can deal with change resistance and ensure that the effort to specify information requirements and to improve information flow is deliberate and strategic. They can focus on addressing issues that impact business performance, project delivery, productivity, and decision quality.

In MindCollect, search and faceted filtering enables users to focus on information requirements associated with specific processes, IT systems, roles, and tasks/decisions. This analytical capability enables users to strategically focus on improving information flows that are key to the success of core business processes. Using simple interfaces, similar to the ones found on Amazon.com, users can see, for example:

- The business objectives, business units, roles, tasks, decisions and KPIs associated with each process
- The information inputs needed by each task and role, and the information outputs produced by each task and role
- The processes, business units, and roles with the greatest number of issues and the top issues for each
- A list of the issues that impact each goal or KPI associated with each process
- The issues that impede the performance of each task
- The most “connected” issues with systemic impact for the most organizational units, processes, roles, tasks, and decisions
- The estimated financial costs and amount of time wasted by each issue
- Recommended solutions for issues and their associated processes, business objectives, KPIs, roles, tasks, and decisions.

APPLICATION CASE STUDY

The State of Vermont's State Emergency Operations Center (SEOC) is responsible for coordinating statewide response to manmade and natural disasters. Response activities involve interaction, cooperation, and coordination with State agencies, local municipalities, utilities, local incident commanders, regional planning organizations, and Federal/other State entities to provide services to the population.

Most EOC's operate with myriad processes involving multiple stakeholders. This makes it difficult to fully understand and generate a common operational picture. Stakeholders that are expected to collaborate effectively during a crisis often do not have a shared understanding of the processes that need to be implemented, the tasks that need to be completed, and the information they need to share with their counterparts. These critical processes, some simultaneous, some sequential, allow EOCs to, among other things:

- Gain an enduring holistic understanding of the spatial and temporal information environments
- Respond to municipal and citizen requirements (food, water, shelter, information) promptly and efficiently
- Request resources from other states or the Federal government
- Respond to first responder tactical requirements
- Respond to media requests for information
- Determine whether the event warrants declaration of disaster

The VT SEOC leadership recognized the need for an efficient and effective way to improve processes, elicit information needs and task details from their staff and external partners, and identify necessary software and IT integration requirements. Mind-Alliance Systems delivered a process mapping workshop for the VT SEOC centered on using MindCollect to elicit structured user stories (task narratives) from domain experts involved in the response to dam failure disasters. Within roughly six hours during a three-day period, 50 state participants, some of whom had never met, used MindCollect to define processes, tasks, and the information needs associated with each task. Stakeholders now had a common understanding of the information they needed to share with each other during a crisis and expressed 96% confidence that they achieved all of the workshop learning objectives. The simple Q&A-based mechanism in MindCollect enabled domain experts to author machine-readable task narratives with information requirements. Without the use of MindCollect, these results would not have been possible in such a short period of time.

VT SEOC used the findings from the workshop to:

- Build process maps for emergency management processes
- Create a central repository of high-quality task data that was needed to produce their key deliverables
- Change/adjust workflows
- Gain access to more/better operational information
- Drive functional requirements of emergency management software
- Improve response times to those in need
- Provide better executive support to government decision makers and elected officials

CONCLUSIONS & FUTURE WORK

The capability to profile information requirements in both a human and machine-readable format is valuable because it captures and formalizes institutional domain knowledge, enabling analysis that clarifies how information flows between functions. Furthermore, it supports informed development of strategies, budgets, and plans for closing information gaps in order to enhance business performance, decision quality, situational awareness, planning, task productivity, and service delivery.

Future research will focus on leveraging semantic knowledge representation, text analytics, and machine-learning technology to express information requirements inside of a user's normal workflow and automatically match them to information resources. Linking ontologies and datasets together using Linked Data technologies will enable diverse externally curated and highly valuable knowledge sources to be made available to users.

With information being generated at such a prodigious rate, the challenges facing knowledge workers are not only finding the right datasets but "sense-making", curating, versioning, maintaining, indexing, searching, querying, retrieving, and re-using information, enabling '**data**' to find '**data**'.

Private corporations and governments are increasingly realizing the potential of the Semantic Web and of encoding knowledge as Resource Description Framework (RDF). RDF [12] is a meta-data model and the language of choice for conceptual description and modelling of information on the World Wide Web. It allows semi-structured data models, also known as ontologies, of domains of interest to be built. Facebook [13] uses Open Graph Protocol (OGP) [14] allowing a Facebook user to integrate other non-Facebook web pages into the user's social graph. OGP uses machine-processable semi-structured data to mark up web pages. Various governments, in order to improve the delivery of services to their citizens, are opening up their data and publishing these data in semi-structured format, many of them in RDF, to improve the delivery of goods and services. The United States government has set up *data.gov* to release public data. The UK Government, keen to unlock the benefits of economic and social gain of public sector information (PSI) reuse, has set up *data.gov.uk*.

The continued adoption and usage of ontologies and semi-structured data in government and industries is bringing about the growth in datasets published in linked data format, and a growing interest in connecting these datasets together. Linked Data [15] is a style of publishing data on the Web that emphasises data reuse and connections between related data sources. With these datasets in different knowledge bases and data stores, there is a paradigm shift occurring. This shift is an important one. We are moving away from the paradigm of "given a set of data, what technique(s) can I use on this dataset and gain insights" to the paradigm of "given a problem, what is the best dataset I can get to solve the problem or answer the questions." The potential to leverage Semantic Web and Linked Data in the field of emergency preparedness planning is significant.

Smart and Shadbolt [16] argued for a Semantic Battlespace Infosphere (SBI), a knowledge infrastructure for coalition interoperability using the Semantic Web and organized around a framework for advanced modes of information integration, exploitation and exchange in coalition military contexts. They showed how an SBI can solve the 'epistemic inter-operability' problem caused by an incompatibility that may exist between coalition partners with respect to conceptual models, inference processes, and reasoning strategies. An SBI can be used to realize an infrastructure with the ability to exploit semantically-enriched representations and establish mappings between disparate entities to integrate and fuse information from semantically heterogeneous information sources.

Such an SBI, applied to emergency preparedness planning, and in the hands of users with the appropriate query tools, can be used to efficiently discover data and expertise, matching the right people to appropriate resources.

MindCollect has developed taxonomies, ontologies and datasets that can aid in profiling information requirements and information-related behaviors in crisis management and other domains. In addition, linking ontologies and datasets together using Linked Data technologies will enable diverse externally-curated and highly valuable knowledge sources to be made available to users.

A future version of MindCollect, called MindPeer, will leverage more domain ontologies, web semantics and predictive models generated with machine-learning techniques to profile various types of information requirements and deliverables. The system will guide users to relevant data in an automated and semi-automated interactive and exploratory manner. Increasingly sophisticated and personalized human-computer interaction techniques will enable users to mount efficient and high quality queries against the semantics-enabled datasets. In certain contexts, such as national intelligence, this capability has the potential to contribute not only to collection and analysis, but to the transformation of Producer/Consumer relations as described by Frank [17].

REFERENCES

1. Serena, C.C. (2014). Lessons Learned from the Afghan Mission Network: Developing a Coalition Contingency Network. RAND Corp.
2. Office of the Assistant Secretary of Defense (OASD) for Network Infrastructure and Integration (NII). (2010). Version Description Document for The DoD Architecture Framework (DoDAF) and DoDAF Meta Model (DM2) Version 2.02. Department of Defense, 30 Sept. 2010.
3. Rodhain, F., Fallery, B. (2008). "What is Information Requirement? A Constructivism versus a Positivism Perspective. in D. Jawahar (Ed). Emerging Issues in Management Research, pp.1-20. New Delhi: EXCEL BOOKS.
4. NAF: NOV-3 Operational Information Requirements Subview http://trak-community.org/ index.php/wiki/ NAF%3ANOV-3_ Operational_Information_Requirements_Subview
5. Kamien, D.G., Cloutier, J-F, Ranger, D. (2005). "Needs Analysis for Information Sharing" in D. Kamien (Ed.)The McGraw-Hill Homeland Security Handbook (1st edition). New York: McGraw-Hill.
6. Cloutier, J.F. & Kamien, D. (2012). "Key Requirements for Enabling Information Sharing Planning and Policy Audits" in D. Kamien (Ed.) The McGraw-Hill Homeland Security Handbook (2nd edition). New York: McGraw-Hill.

7. United States Patent 8,738,325 “Method and system for analyzing information transfer among a plurality of parties” May 27, 2014.
8. Cloutier, J-F, Kamien, D., & Desourdis, R.I. “Channels: An Information Flow Modeling System to Support Planning and Interoperability” in Technologies for Homeland Security (HST), 2011 IEEE International Conference, November 15-17, 2011.
9. NATO Advanced Research Workshop (ARW) on "Building Inter-Agency Information Sharing Plans for Homeland Security,” Brussels April 4-6, 2011. Funded by NATO Science for Peace and Security Program.
10. http://swizec.com/blog/natural-language-generation-system-architectures/swizec/4535
11. Dale, R. & Reiter, E. (2000). Building natural language generation systems. Cambridge, U.K.: Cambridge University Press.
12. Resource Description Framework at: http://www.w3.org/RDF/ .
13. Facebook at https://www.facebook.com/ .
14. Open Graph Protocol at: http://ogp.me/ .
15. Linked Data at: http://linkeddata.org/ .
16. Smart, P.R. & Shadbolt, N.lR. (2007) The Semantic Battlespace Infosphere: A Knowledge Infrastructure for Improved Coalition Inter-Operability. In, 4th International Conference on Knowledge Systems for Coalition Operations (KSCO), Waltham, Massachusetts, USA, 01 - 02 May 2007.
17. Frank, A.B. (2015) “Transforming Produce/Consumer Relations through Modeling and Computation,” in R. Arcos and R. H. Pherson, (Eds.) Intelligence Communication in the Digital Era: Transforming Security, Defense and Business. Basingstoke, UK: Palgrave Macmillan.