Adaptive Integration of Information Supporting Decision Making: A Case on Humanitarian Logistic

Kelli de Faria Cordeiro
Center for Naval Systems Analysis of Brazilian Navy
Federal University of Rio de Janeiro
kelli@ppgi.ufrj.br

Maria Luiza Machado Campos
Federal University of Rio de Janeiro
mluiza@ppgi.ufrj.br

Marcos Roberto da Silva Borges
Federal University of Rio de Janeiro
mborges@ppgi.ufrj.br

ABSTRACT

There is an urgent demand for information systems to gather heterogeneous information about needs, donations and warehouse stocks to provide an integrated view for decision making in humanitarian logistics. The dynamic flow of information, due to the unpredicted events, requires adaptive features. The traditional relational data model is not suitable due to its schema rigidity. As an alternative, Graph Data models complemented by semantic representations, like Linked Open Data on the Web, can be used. Based on both, this research proposes an approach for the adaptive integration of information and an associated architecture. An application example is discussed in a real scenario where relief goods are managed through a dynamic and multi-perspective view.

Keywords

Knowledge Integration, Humanitarian Logistics, Graph Databases, Decision Support System, Linked Open Data

INTRODUCTION

The decision making process, in complex environments, faces several challenges due to the dynamic of unpredicted events. An example of this kind of scenario is the humanitarian logistic, as we could see in recent disaster caused by the Typhoon Haiyan in the Philippines. The amount, type and location of needs and donations are unpredictable. In the assessment phase, it is crucial to evaluate the victims and staff needs to better support the following phases (Widera et al., 2013). Such task is accomplished during high dynamic events, such as the request of unpredicted resources, the donation of unsolicited relief goods and the availability of warehouse stock. The difficulties can increase with requests of special needs and insufficient available resource. To manage this complex environment, some researches focused on operational models (Franke et al., 2011, Münzberg et al., 2013) and others on the design of operational process (Widera et al., 2013).

However, from information management perspective, some particular characteristics emerge. There are many different organizations and systems generating and consuming information with semantic and structure heterogeneity. In this scenario, to deliver an integrated view of the situation requires adaptive procedures where information plays a central role. Some information systems are available to put these information together, e.g., Sahana Eden. However, traditional information systems based on relational database models are not suitable to address requirements of information with unpredicted structure and insufficient descriptions.

Focusing on the presented issues, this paper proposes the design of an architecture to support the adaptive integration of information with semantic and structure heterogeneity. The approach uses graph representation and Linked Open Data (LOD) principles as a strategy for data structure and description adaptation. Moreover, the architecture uses available worldwide information on the Web to address unexpected demands for complementary information. With these features, it is possible to provide an integrated view according to the user's role. The managers, citizens and community can then have integrated information to make decisions as volunteers, donors and victims, according to the situation. The proposed approach was applied in a real scenario and illustrates how information integration can be adapted to an ongoing situation.

MOTIVATION SCENARIO

The scenario case characteristics were elicited based on reports about response phase logistics cases (Howden, 2009, Widera et al., 2013). But, one scenario case was chosen to discuss the application of the proposed approach. Despite the local disaster scope, this case has examples of information dynamic and its semantic heterogeneity to illustrate the integration problems that occur in national or continental disasters as well.

In January 2013, a nightclub got fire in Santa Maria City, in the south of Brazil. Based on official police report, there were about one thousand people when a rock band's pyrotechnics equipment put fire on the roof, and the flames spread very fast. The roof was made of acoustic foam which releases a toxic gas, called cyanide, when on fire. Over two hundred people died that night, and over one hundred suffered severe pneumonitis. By the eighth day, the medication dispensed to over one hundred victims was ineffective due to the lung damages of cyanide gas. As an alternative treatment, an experimental procedure was offered. The extracorporeal membrane oxygenation uses the ECMO equipment to support oxygenation to patients with lung damages. In addition to other regular needs, this new alternative treatment had heavy demands. There was a request for one hundred ECMOs equipment, as well as to associated support equipment, such as ventilators. At the same time, a German company offered the donation of the ECMOs. As a result, the responder managers had to familiarize themselves with the new technique and equipment.

AN APPROACH FOR ADAPTIVE INTEGRATION OF INFORMATION

In the last years, some non-relational data representation models have been used to support dynamic information requirements. One of them is based on graph structure. In graph databases, there is no semantic structuring schema. Data can be stored without identification of the entity to which they belong. The identification can be done further through creation of new edges and vertices (Robinson et al., 2013). In parallel, a World Wide Web (W3C) initiative, called Linked Open Data (LOD), was proposed (http://w3.org/DesignIssues/LinkedData). It is based on Resource Description Framework (RDF), a directed labelled graph format for representing information on the Web. The potential virtues of LOD, explored in our solution, are: (i) the possibility of building the data schema at a later time, (ii) the entity identification being determined by the context of use, (iii) the integration and reuse of data on demand, and (iv) the co-existence of heterogeneous semantic structure in the same dataset.

A Graph Data or Property Graph is a directed graph which is composed of vertices and directed edges described by properties representing entities and relationships (see Figures 1a and 1b). The simplicity of this representation brings flexibility to its use in different situations. Figure 1c illustrates an example of our scenario location information. The entity *Santa Maria (Brazil)* is described by the entity *Place* and is a subdivision of the *South Region Brazil* entity. The vertices and directed edge of Graph Data can be implemented using RDF, which consists of a set of triple statements in the `<subject, predicate, object>` form (see Figure 1d). The subject identifies the resource to which the statement refers. The resource is identified by a Uniform Resource Identifier (URI), composed of a Uniform Resource Locator (URL) and a name. The predicate corresponds to a relationship or a property of this resource. And the object represents a value or a resource associated with the subject through a predicate. The Figure 1e depicts the RDF graph example of Figure 1c with its corresponding URIs. The RDF statements can be serialized as RDF/XML code, as illustrated in Figure 1f.

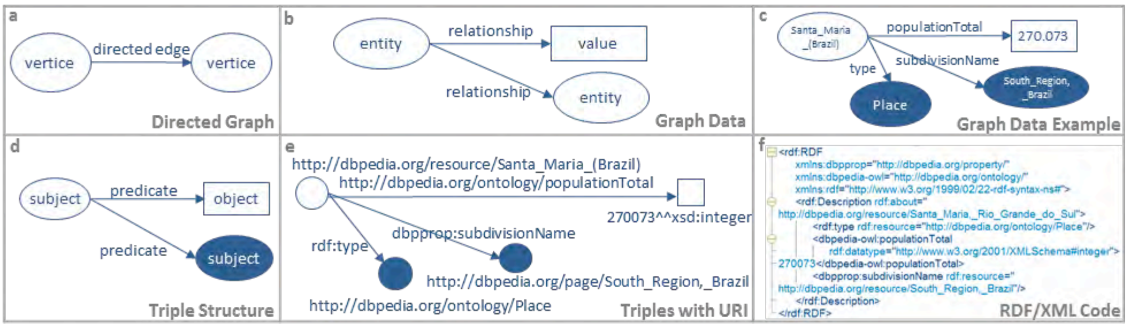

Figure 1: Graph Data and Triple Structure with Examples

The LOD Web, or Web of Data, is a powerful environment, which can be used as a worldwide source of information with billions of triples published in hundreds of datasets. The Web of Data functional characteristics are suitable for contextual adaptation systems which must process information even without prior knowledge of its structure. The Web of Data has been used and discussed on researches about Emergency Management Information Systems (Cordeiro et al., 2011, Schulz et al., 2012). These researches have demonstrated that linking features of LOD address many problems in dynamic environments.

The links on the Web of Data are primarily intended to aggregate value to data. They can be used to describe, associate and infer new knowledge, besides supporting navigation and exploration. Descriptions can contextualize the data, improve the semantics, add new information and reveal the source of the data. The interlink can be used as an important semantic enrichment mechanism: (i) to describe information resources through an association to other data resource that corresponds to a new attribute; (ii) to associate information using new interlinks on triple, graph or dataset level; and (iii) to discover or infer new information or connections by navigating through the links from one resource to another. Based on these functions, the integration is performed when different data sources are interlinked by semantic relations, when data is interlinked to conceptual reference frames which are also published in LOD format, and through the use of vocabularies and ontologies widely used. The following section explores these functions to support adaptive integration of information according to user's role and situation.

ARCHITECTURE FOR ADAPTIVE INFORMATION OF HUMANITARIAN LOGISTIC

Based on Graph Data adaptive features and Web of Data interlinking main functions, the proposed architecture is composed of four layers presented in Figure 2. The top one is the dynamic environment where unpredicted events occur. Several agents, with different roles, interact with each other and unexpected events may happen. In such complex interactions, information is exchanged in different ways. As a result, the information management needs to adapt to the evolving situation and is supported by following layers.

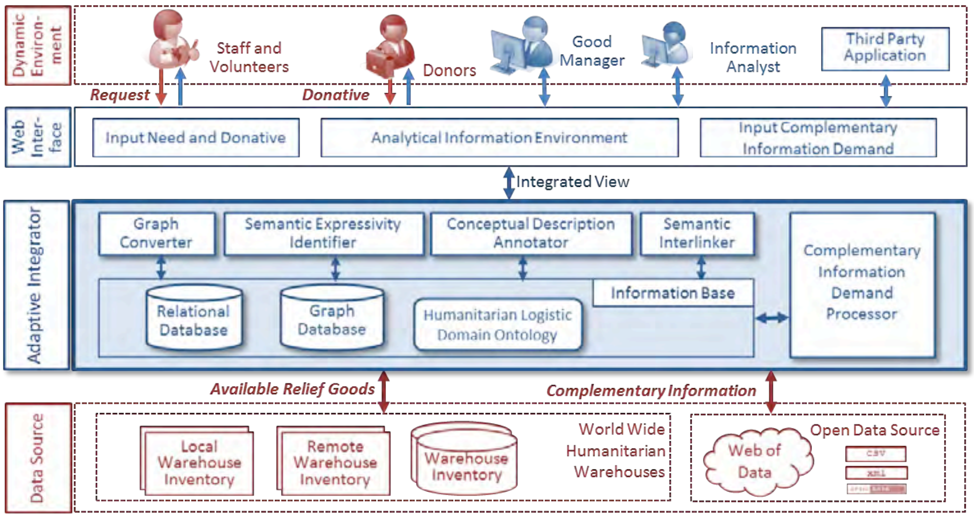

Figure 2: Architecture for Adaptive Integration of Humanitarian Needs and Donatives Information

The Web Interface layer has components that allow the input of heterogeneous information, analysis of integrated information and the request of complementary information. All communications are made using Web standards (http, ftp) and structured data file formats (csv, xls, xml, dbf). Furthermore, third party analytical applications can be built upon this layer.

The middle layer is the Adaptive Integrator. Its components are responsible for integrating information with unpredicted structure from different sources and semantic expressivity levels. These components are orchestrated by ETL4LOD (http://greco.ppgi.ufrj.br/lodbr/index.php/principal/etl4lod-2/) which implements data workflows. The Graph Data Converter transforms the structure of the input data into a graph, encoded in RDF. The Semantic Expressivity Identifier analyses the RDF graph and assesses the level of the data descriptor. A resource described by a property and a value belongs to level 1. A resource described by a property and other resource belongs to level 2. A resource or a property described by a vocabulary or an ontology, belongs to level 3. With the level assessed, the Conceptual Description Annotator augments the semantic expressivity of the resource descriptor using conceptual reference models. For example, a resource level 1 can be interlinked to other resource which represents an ontology class. As a result, the resource moves from level 1 to level 2. With the semantically enriched data, the Semantic Interlinker creates new links among converted and semantic evolved RDF graph enabling the adaptive integration of heterogeneous information. New links are created extending tools like Silk (http://wifo5-03.informatik.uni-mannheim.de/bizer/silk/). The last component is the Complementary Information Demand Processor, which crawls the open data cloud to retrieve additional information adapting to meet unexpected demands.

All processed data is stored in the Information Base sub layer, where there are two types of databases, with and without schema. The Relational Database stores previous and legacy information, e.g., warehouses' inventory. The Graph Database stores current data about requests and donations of relief goods using RDF representation. The processed and interlinked RDF data are also stored in this database using a triplestore, such as, Virtuoso (http://virtuoso.openlinksw.com/). Furthermore, the Humanitarian Logistic Domain Ontology is stored in this layer. The lowest layer comprises the data sources of worldwide warehouse's inventories and complementary information. The first one can be local or remote. The second one is composed of remote data sets available on the Internet. It can be the Web of Data Cloud or any other structured open source data.

Architecture Application Prototype

The proposed architecture was applied considering the dynamic of the motivation scenario. The incoming files used the Web Interface to feed the system. The request of ten doctors specialized in pulmonology came in a csv file. The need of fifty ECMO and the associated medical ventilator was written in an xls file. The donation of the ECMO equipment by a German Company arrived in xml file. And the stock inventory was stored in a relational database table. The Graph Converter Module transformed these heterogeneous data to RDF format. The graph representation of the converted data is illustrated within dashed rectangles in Figure 3.

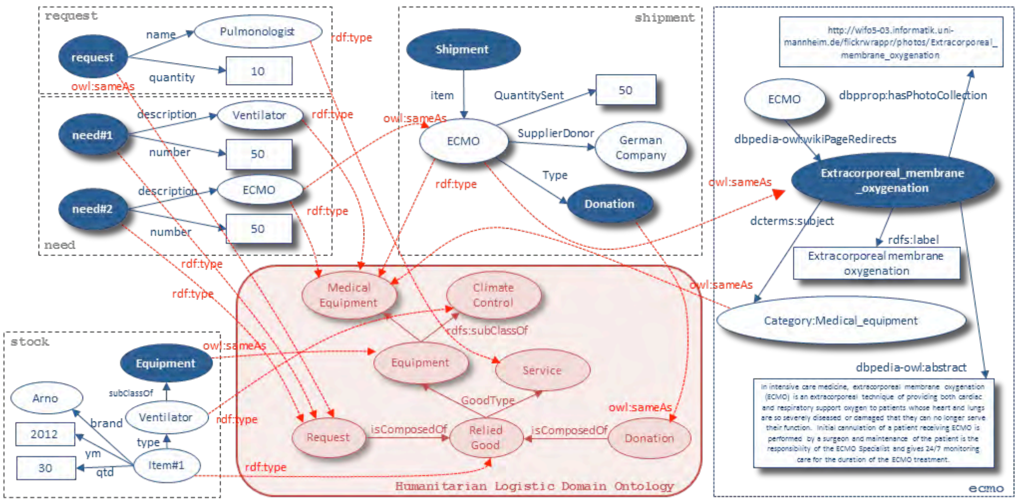

Figure 3: Graph representation of the scenario case integrated view

The Semantic Expressivity Identifier recognized lower level on the converted RDF data, as some resources does not have the associated entity. To semantically enrich these information resources, the Conceptual Descriptor Annotator created new links to resources which represent entities of the domain ontology. For example, the ECMO and Ventilator information resource was described as medical equipment by linking them to the *Medical Equipment* entity. In contrast, the stocked Ventilator was described as *Climate Control* equipment, making explicit that it is a fan. By inference, these resources are classified as a *Relief Good* entity. In the scenario, there was a demand for additional information about ECMO. To address that, the Complementary Information Demand Processor retrieves information about the ECMO therapy and equipment from the Web of Data. Having all information resources described by common entities, the Semantic Interlinker created new links mapping common concepts, allowing the design of an integrated view. The new links are represented by dashed arrows in Figure 3 and the information about ECMO is represented in graph format within dotted rectangle.

With this approach, relief goods management can be supported by a dynamic and multi-perspective view. For example, from the warehouse manager's perspective, it was found out that there are no ventilators in stock, only fan for climate control. Another example is from the donors and medical staff's perspective view. There were ECMOs available, came from donation shipment, but they were not enough. Besides, the hospital crew was not familiarized with ECMO therapy. To illustrate the results of the whole process, Figure 4 depicts the layers behind the analytical information system interface which supports medical staff and donors. The first one illustrates the graph representation of the information base interlinked to the dataset published on the Web of Data. The second illustrates the SPARQL query code and its results. And the third layer is the user interface with information about the ECMO. Based on queries results, some decisions were made. Managers sent an acquisition order of medical ventilators, donors prepared a shipment with 10 more ECMO equipment, and the hospital crew could get familiarized with the ECMO therapy, studying some photos with operation procedures.

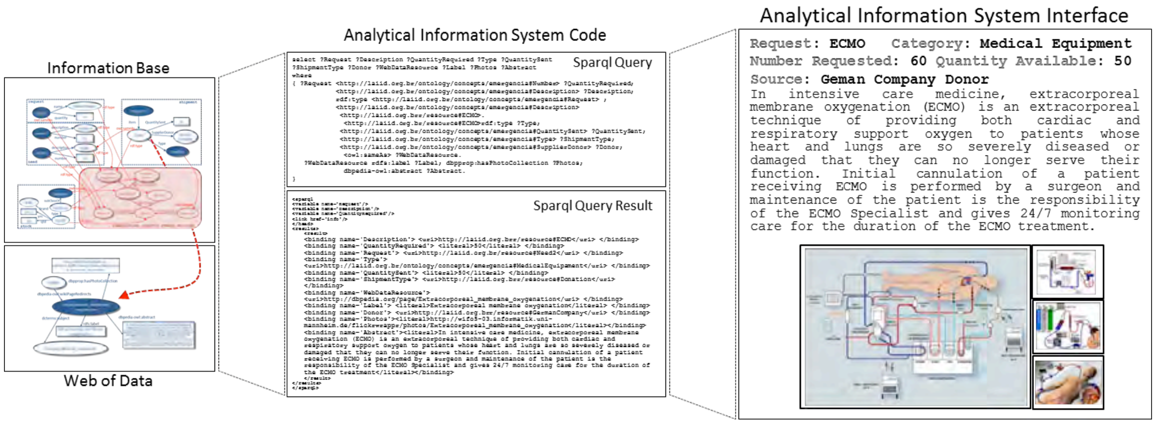

Figure 4: SPARQL query result example and representation of integrated information under adaptive approach

CONCLUSION

Using Semantic Web standards and Graph Data representation, exploring data available on the Web and extending existing tools it is possible to support the adaptation of information bases to the dynamic behaviour of unpredicted events such as those present in humanitarian logistics. The research results suggest that even though the incoming files have different descriptive structures and different semantic expressivity levels, an integrated view can be built to support decision making in the assessment phase. Mode details are available at http://greco.ppgi.ufrj.br/lodbr/index.php/principal/laiid. Keeping on real scenario cases, the next application prototype will be developed using information from logistic systems used to support disasters, such as Sahana Eden deployment to support Typhoon Haiyan in Philippines (http://eden.dswd.gov.ph/eden). In a preliminary analysis, there are 31% of records with text free comments and 32% of blank cells, indicating and reinforcing the need of adaptive semantic structure in this domain. Besides, graphical friendly user interfaces will be used.

REFERENCES

1. Cordeiro, K.F.; Marino, T., Campos, M.L.M.; Borges, M.R.S. (2011) Use of Linked Data in the Design of Information Infrastructure for Collaborative Emergency Management System. International Conference on Computer Supported Cooperative Work in Design (CSCWD). Swiss.
2. Franke, J.; Widera, A.; Charoy, F.; Hellingrath, B.; Ulmer, C. (2011) Reference Process Models and Systems for Inter-Organizational Ad-Hoc Coordination - Supply Chain Management in Humanitarian Operations. Proceedings of the 8th International ISCRAM Conference. Portugal.
3. Howden, M. (2009) How Humanitarian Logistics Information Systems Can Improve Humanitarian Supply Chains - A View from the Field. The 6th International ISCRAM Conference. Sweden.
4. Münzberg, T.; Berbner, U.; Comes, T.; Friedrich, H.; Gross, W.; Pfohl, H.-C. & Schultmann, F. (2013) Decision Support for Critical Infrastructure Disruptions: An Integrated Approach to Secure Food Supply. In: Proceedings of 10th International ISCRAM Conference. Germany.
5. Robinson, I.; Webber, J. and Eifrem, E. (2013) Graph Databases. Neo Technology, Inc. O`Reilly Media, Inc.1th Edition. ISBN-13: 978-1449356262
6. Schulz, A.; Paulheim, H.; Probst, F. (2012) Crisis Information Management in the Web 3.0 Age. In Proceedings of the 9th International ISCRAM Conference. Canada.
7. Widera, A.; Dietrich, H.-A.; Hellingrath, B.; Becker, J. (2013) Understanding Humanitarian Supply Chains - Developing an Integrated Process Analysis Toolkit. In 10th International ISCRAM Conference. Germany.