COBACORE

Community Based Comprehensive Recovery

Grant Agreement № 313308

D2.2 Definition of Modelling Hierarchy

WP2 Information Framework

Version:	1.0
Due Date:	31st March 2014
Delivery Date:	31st March 2014
Nature:	R
Dissemination Level:	PU
Lead partner:	University of Ulster
Authors:	GeoPii, Integrasys, Tilburg

www.COBACORE.eu

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 313308.

Deliverable Title Definition of Modelling Hierarchy

Deliverable Number D2.2

Keywords: Data Framework, Crisis Management, Ontologies, Semantic Intelligence, Interoperability, Data Standardisation, Data Governance.

Executive Summary

Disasters and large scale emergency incidents present a number of distinct information management challenges. Specifically, data needs differ across stakeholder groupings depending on roles and responsibilities, data needs change across the disaster cycle; whilst the dynamic nature of the environment is not conducive to data sharing. To date the emphasis has been placed on the barriers and obstacles to data sharing amongst professional responders and how inter and intra operability can be enhanced amongst the professional responder community. In contrast, limited emphasis has been afforded to the information needs of the affected community, including the means by which communication flows between professional responders and the affected community could be bolstered in order to improve situational awareness. Moreover, greater emphasis needs to be placed on information sharing between professional responders and volunteer responders from within the affected community, in order to more effectively harness community capacity and ensure effective co-ordination of this vastly under-utilised resource base.

The COBACORE platform is intended to facilitate enhanced communication flows between the professional responder community, the affected community and volunteer responders in order to enhance situational awareness, inform and guide response planning and ensure more effective co-ordination of the activities of the volunteer responder community across different points in the disaster cycle. To facilitate such outcomes it is necessary for the COBACORE data framework being developed in Wp2 will need to extend beyond conventional information systems in crisis management. The 'additionality' in data provision is where COBACORE is considered to enhanced value to the crisis response process. However, the objective of COBACORE is not to replace existing disaster management information systems but to compliment and extend their analytical capacities and to enhance stakeholder collaboration.

Achieving such an objective does however present a number of design and feature challenges for COBACORE pertaining to interoperability. Deliverable 2.2 presents an architectural overview of the COBACORE data framework to include detailed understanding of the design features and functional requirements for effective operation, as well as the crisis response ontologies, vocabularies and semantic modeling required in order to optimize interoperability and enhance usability from the perspective of both professional responders and affected communities.

Table of Contents

.

1.0 INTRODUCTION

There is general consensus across multi-agency groupings affiliated with disaster response and recovery planning of the inherent value derived from timely, robust and credible data as a means of informing decision making across all phases of the disaster cycle. Indeed, the absence of key datasets which can culminate in an ‘information vacuum’ serves to undermine the effectiveness of intervention measures adopted during the acute response phase of the disaster cycle as well as jeopardising the sustainability of medium-long term recovery and reconstruction programmes. Deliverable 2.1 provided a detailed overview of the types, formats and potential sources of data considered integral to informing crisis response planning within a series of crisis management domains comprising, institutional and government, mobility and transport, vital and critical infrastructure, social, cultural and educational, healthcare, economic, environmental as well as safety and security.

Whilst it is apparent that different forms of crisis require bespoke data in order to initiate an effective and co-ordinated response it is also apparent that there is considerable overlap in terms of the baseline data needs and in particular the static indicators such as population demographics, socio-economic profile etc. which are equally important in serving to compliment dynamic data such as the number of injuries/fatalities, potential cascading effects etc. Nonetheless, such essential baseline indictors are not always readily available or in a format conducive to interstation and analysis. Consequentially, whilst there has been marked improvements in data collection and analysis in the last decade resulting in more complete global databases on disaster impacts, leading to better risk assessment and more effective targeting of prevention and preparedness activities (UNFPA, 2010) it is also apparent that data sharing capacities have not presently reached the levels of sophistication to facilitate full exploitation of advancements in ICT services.

There are a number of reasons for the deficiencies in data sharing. Disaster planning and recovery is by nature multi-agency and is often cross-jurisdictional. It is characterised by the existence of ‘operational silos’ which can be exacerbated by heterogeneous organisational structures and information frameworks which are very often not geared towards collaborative working. Political resistance (both within and between organizations and at governmental level) has been cited as a further barrier to information sharing. Moreover, the nature of the working environment, particularly within the acute response phase, is highly dynamic, necessitating almost instantaneous decision making. These timelines generally do not lend themselves to data/information sharing unless protocols are already in place. In addition there is also recognition that responders are more receptive to *receiving* information specific to their requirements and are not as highly incentivised with *providing* information to others who might directly benefit. Research has shown that developing a greater understanding and appreciation of the work processes and data needs of different stakeholders involved across the crisis response planning process can have hugely positive impacts on information sharing and co-ordination of response as well as guiding the future

development of disaster management information systems in order to enhance their usability and interoperability (Bharosa et al, 2009).

Existing research has for the most part nonetheless been heavily focussed on improving the effectiveness and efficiency of data and information sharing amongst professional responders. Much less attention has been afforded to the deficiencies in the information flow between the affected community and professional responders and between the professional responders and volunteer responders. The over-arching objective of COBACORE is therefore to enhance the flow of communication between the professional responder community, the affected community and volunteer responders across all phases of the disaster cycle. The 'enhancement' of communication flows is considered to encompass the timeliness of data/information exchange, the accuracy and comprehensiveness of data exchange, the frequency of exchange between the aforementioned key stakeholder groupings as well as utilising advancements in ICT and social media in order to extend the 'reach' of communication flows in order to embrace as wide an audience as possible.

1.1 Deliverable Context

The COBACORE data framework is not to intended to replicate or replace existing disaster information systems but to complement and enhance the existing Disaster Management Information Systems (*DMIS)* offer through the development of a platform which is both value enhancing and user-friendly. COBACORE will serve as information conduit which can interface with existing crisis management information systems to afford more robust and comprehensive data profiling. Pre-loading' key datasets which predicate crisis response planning into one centralised accessible repository will permit more robust and sophisticated analytical interpretation culminating in improved situational awareness and a more credible evidence base upon which to inform and guide crisis response planning. The COBACORE data framework is also intended to more effectively integrate and interpolate dynamic data capture and in particular crowd-sourced data such as Twitter, Facebook etc.

It is widely recognised that the potential of crowd sourced data within the confines of data response has not as yet been fully exploited. Indeed the utilisation of social media thus far has been predicated on information gathering (on the part of professional responders). COBACORE is intended to further develop the utilisation and application of social media platform both as a credible data source (pertaining to data validation and verification) but also as a real time data distribution medium which enables professional responders to disseminate information to affected communities.
The defining added value feature of the COBACORE data framework and indeed the feature that in essence sets it apart from existing disaster management information systems is however the propensity to not only capture needs and capacities across affected communities but through the application of semantic intelligence modelling to 'match' needs and capacities culminating in activity generation (Figure 1). The extent and depth of the 'matching' exercise will be further refined as we move

towards final data framework development and in conformance with the functions and feature specifications emanating from WP3.

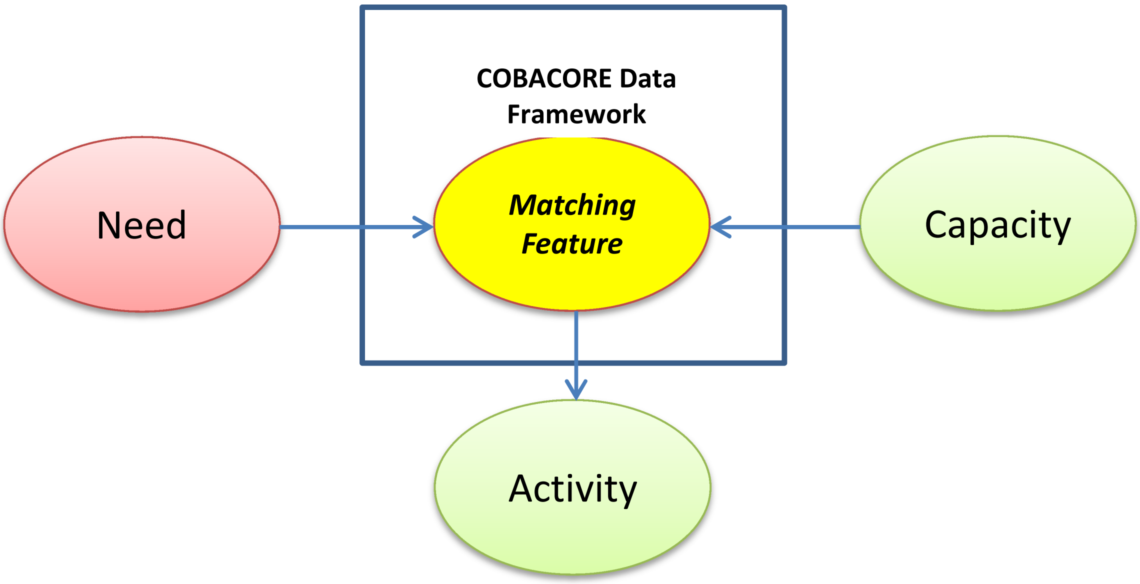


Figure 1: COBACORE Data Framework Needs Matching Feature

To achieve the stated objectives considered thought must be given to the platform design in order to maximise inter-operability (relative to existing DMISs and across jurisdictions) and transferability (across disaster types). In the immediate aftermath of a crisis, needs have to be identified and prioritised before the emergency actions are taken, so as to provide proper humanitarian assistance in a timely and co-ordinated manner. Decision makers are nonetheless often faced with the challenge of developing a response strategy predicated on a wide variety of data sources which may appear at once and in unstructured format which is not conducive to automated (machine processed) interpretation. Moreover, the unstructured information may also include different expressions for similar needs. For example, some people may use different food brand names to express their food needs. Ontologies are used within crisis response to provide structured information which creates meaningful relationships between information resources and to allow machines to process, infer, or combine the information from different sources automatically into a consistent body of knowledge. This culminates in synergies in the disposition of data resources whilst enabling time-critical response activities to be accelerated.

1.2 Positioning of D2.2 within COBACORE Project Structure

Work Package 2 seeks to develop a disaster management data framework for the operationalization of the COBACORE platform being developed in Work Package 4. The data framework will be developed in conjunction with WP1 in order to validate key data requirements for advancing crisis decision making and policy support.

Collaboration with WP1 will also enable gaps and deficiencies in the existing information exchange process to be identified affording insight to the necessary focus and positioning of the COBACORE data framework in order to optimise impact and value enhancement.

Building upon the data identification (D2.1) this deliverable also includes input from the (underlying discussions and work of) deliverables of other work packages. In particular D1.1 has a strong –mutual– connection with the following deliverables.

- **D1.1 Scope and Requirements, Domain Analysis**
  Deliverable 1.1 outlines scope and requirements of the COBACORE concept based on the domain analysis. From this domain analysis specific attention was given to the involved actors in the disaster response operations. While many different persons and stakeholders can be identified, they can be classified in three major groups, based on their background and relation to the disaster comprising, professional responders, volunteer responders and affected community.

- **D3.1/MS3.1 Functional behavior**
  From the issues and functions presented in this document we distill the various functions that the platform is required to provide in order to address these issues and take leverage the identified potential. The deliverables from work package 3 describe in more detail how these functions work.

- **D4.1 Platform requirements**
  The platform requirements outline the capabilities of the platform in order to provide the describe functions and details. This includes for example the underpinning technical infrastructure and the abilities the platform should provide for example in terms of user interface and interactions.

- **MS4.1 / MS 5.1 Intermediate evaluation & Performance indicators**
  The outcome of the domain analysis and the resulting opportunities and issues are used by Work package 5 in their deliverables to base the indicators and evaluation setup upon. Vice versa, the outcomes of the evaluations undertaken by Work Package help to assess the validity of the outcomes of Work Package described in this deliverable. The functional needs and specification requirements are considered within this deliverable.

2.0 DEFINED DATA FRAMEWORKS FOR DISASTER RESPONSE

The identification, management, synthesis and analysis of spatial and temporal data are central components of post-crisis needs assessment for reconstruction and recovery globally. Indeed, preparedness for and response to crises is dependent on the availability, accessibility and usability of data. Data must be transferable and communicable to multi-agency partners co-ordinating the crisis response and to citizens affected by the disaster. However, to achieve this interoperability (across professional responders) and between the professional responders and affected communities, there needs to be a common approach taken in order to ensure an effective and efficient response, reconstruction and recovery model. This common approach culminates in enhanced interoperability of disaster management information systems. Interoperability within the confines of crisis response generally comes in three aspects: (1) Governance Conformance: practices and procedures technical (2) Compatibility of message formats (3) Semantic: terminology and definitions.

The COBACORE data framework architecture is predicated on spatial orientation and utilises location as the consistent variable within the confines of a disaster event which can serve to link and derive commonality across diverse and disparate datasets in order to enhance the functionality and analytical capacity of the COBACORE data framework. As a consequence of the GIS based architecture being developed within the COBACORE framework, Geospatial data standards will govern the design and development. The implications and necessary standards conformance are discussed in Section 2.1. The remainder of the chapter will explore existing ontologies (Section 2.2.) and vocabularies (Section 2.3) which have been applied within the confines of crisis response in order to optimise commonality of understanding and systems integration. The final section details the semantic modelling requirements for the data framework in order to optimise interoperability.

2.1 Data Governance and Standards Conformance

In the context of spatial data and Geographic Information Systems (GIS) as a whole, standardisation is set by the Open Geospatial Consortium (OGC) who are responsible for ensuring that end user needs are addressed by cooperation among data and GIS vendors. The OGC (2005) suggests that these standards are centred upon *“geo processing, which includes capabilities now found in geographic information systems (GIS) and digital systems for Earth imaging, web mapping, location based services, surveying and mapping, CAD-based facilities management, webs of geo located sensors, navigation, cartography, automated mapping etc. The "standards" are consensus-derived specifications for open interfaces, protocols, schemas etc. that enable different vendors' systems to exchange data and instructions, and that enable full integration of these capabilities into all kinds of information systems”*.

The standards set by the OGC are documented through technical specifications that detail interfaces and/or encodings. These are then used by software developers to

engineer open interfaces and encodings in to the architecture of their products. In turn, this addresses interoperability issues which may exist if standards are not applied or adopted, especially when multiple software engineers work independently on products and online services. Through adopting and applying these standards, particularly when more than one software engineer is involved independently, components should essentially 'plug and play' and reduce the need for further debugging (OGC, 2014).

In the context of the COBACORE project, the software development of the GIS based component utilises standardised functionality and data as set by the OGC. This includes the application of, but not exhaustive to, GeoSPARQL, WMS, WFS, WCS, KML, Styled Layer Descriptor, Coordinate Transformation, GeoRSS, GML and GeoAPI. A description of each is presented in Table 1. The premise of ensuring conformity to these standards is to optimise the usability and interoperability of the COBACORE data framework. From an end-user perspective awareness and appreciation of the standards will be less know however non-conformance will severely impinge upon both the interoperability and acceptance (amongst prospective end-users) of the COBACORE concept. The objective of COABCORE is to enhance rather than replace the existing disaster management information systems offer – conformance with the OGC standards facilitates legacy/heterogeneous information frameworks to essentially 'plug' into the COBACORE data framework in order to enhance situational awareness as a consequence of enhanced data provision and analytical capacity.

Name	Description
GeoSPARQL	This standard defines a set of SPARQL extension functions [W3C SPARQL], a set of RIF rules [W3C RIF Core], and a core RDF/OWL vocabulary for geographic information based on the General Feature Model, Simple Features [ISO 19125-1], Feature Geometry and SQL MM.
Web Feature Service (WFS)	Web Feature Service (WFS) represents a change in the way geographic information is created, modified and exchanged on the Internet. Rather than sharing geographic information at the file level using File Transfer Protocol (FTP), for example, the WFS offers direct fine-grained access to geographic information at the feature and feature property level. Web feature services allow clients to only retrieve or modify the data they are seeking, rather than retrieving a file that contains the data they are seeking and possibly much more. That data can then be used for a wide variety of purposes, including purposes other than their producers' intended ones.
Web Mapping Service	Web Map Service Interface Standard (WMS) provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. A WMS request defines the geographic layer(s) and area of interest to be processed. The response

	to the request is one or more geo-registered map images (returned as JPEG, PNG, etc.) that can be displayed in a browser application. The interface also supports the ability to specify whether the returned images should be transparent so that layers from multiple servers can be combined or not.
Web Coverage Service (WCS)	Web Coverage Service (WCS) supports electronic retrieval of geospatial data as coverages – that is, digital geospatial information representing space/time-varying phenomena.
KML	KML is an XML language focused on geographic visualization, including annotation of maps and images. Geographic visualization includes not only the presentation of graphical data on the globe, but also the control of the user's navigation in the sense of where to go and where to look.
Styled Layer Descriptor (SLD)	Styled Layer Descriptor defines an encoding that extends the WMS standard to allow user-defined symbolization and coloring of geographic feature and coverage data. LD addresses the need for users and software to be able to control the visual portrayal of the geospatial data. The ability to define styling rules requires a styling language that the client and server can both understand.
Coordinate Transformation	Coordinate Transformation Service Standard (CT) provides a standard way for software to specify and access coordinate transformation services for use on specified spatial data. This standard addresses a key requirement for overlaying views of geodata from diverse sources: the ability to perform coordinate transformation in such a way that all spatial data are defined relative to the same spatial reference system.
Geographic Markup Language (GML)	The Geography Markup Language is an XML grammar for expressing geographical features. GML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. As with most XML based grammars, there are two parts to the grammar – the schema that describes the document and the instance document that contains the actual data.
GeoAPI	The GeoAPI Implementation Standard defines, through the GeoAPI library, a Java language application programming interface (API) including a set of types and methods which can be used for the manipulation of geographic information structured following the specifications adopted by the Technical Committee 211 of the International Organization for Standardization (ISO) and by the Open Geospatial Consortium (OGC). This standard standardizes the informatics contract between the client code which manipulates

	normalized data structures of geographic information based on the published API and the library code able both to instantiate and operate on these data structures according to the rules required by the published API and by the ISO and OGC standards.

Table 1: Description of OGC Standards used in the COBACORE project

Source: http://www.opengeospatial.org/standards/ct accessed on 6/3/2014

At the European level, spatial data standardisation has been at the forefront of European Directives in recent years. Indeed, most of this attention has stemmed from the INSPIRE Directive, a specification requiring the creation of a cross boundary, pan-European Union Spatial Data Infrastructure (SDI) to facilitate the sharing of GI among all spectrums of government and improving accessibility of GI to European citizens. Under INSPIRE, the 27 nations of the EU are required to undertake a number of changes to their current data capturing methodology to ensure consistency, accountability and comparability of approach . These requirements include:

A. Make available and interoperable, metadata that details the underpinning elements of spatial datasets and services so that they can be shared

B. Implement network services to facilitate an enhancement of awareness, access and use of spatially enabled data

C. The removal of barriers that have in the past, prevented the sharing of GI between different public sector bodies

Pertinently, one of the underpinning drivers behind INSPIRE was to enhance the depth and detail of information available for crisis situations and to promote enhanced levels of standardisation in order to facilitate more effective data transfer both between the professional responding community and across jurisdictions in order to bolster resilience and transpose best practice pertaining to risk analysis and mitigation, response planning and recovery and reconstruction. Organisation such as MapAction have played a pivotal role in show casing the potential of GIS as an analytical tool to inform and guide intervention measures and to visually display the impacts of disaster incidents. The enhanced accessibility of GI has meant that volunteers can remotely make meaningful and valued contributions to inform disaster recovery in terms of data validation and verification.

2.2 Disaster Management Ontologies

The principal means of redressing the issue of interoperability and to facilitate enhanced data sharing across systems is to utilize ontologies[1]. Ontologies provide a unified explanation of concepts and relationships used by the application field, make them interpretable from an IT perspective as well as shareable across different end-users. From a graphic point of view, an ontology can be represented with a graph where nodes correspond to concepts and where links are relations between these concepts.

A variety of application specific ontologies have been developed for disaster information management, which enable interoperability in specific scenarios. Moreover, efforts are already underway to enhance the capacity for data sharing and analytical capacities by combing existing ontologies with geospatial semantic web technologies for emergency response and disaster management (see Bakillah et al., 2007; Lan et al., 2009; Zhang, 2007). In spite of the progress that has been made in recent years the utilisation and application of crisis management ontologies remains relatively immature (relative to other industries/sectors). In order to make meaningful inferences on the added value and interoperability enhancement that could be derived from ontology application, it is first necessary to appreciate and recognise the present state of the art pertaining to disaster management ontology application.

Liu et al. (2013) conducted a review of state of the art ontology design and usability for crisis management. As an initial phase of the ontology review process Liu et al compiled a review of contemporary crisis information systems. This review highlighted that 26 ontologies presently exist that are relevant or could be implemented within the confines of crisis management. The ontologies comprise a combination of ontologies created specifically for disaster management as well ontologies designed for other domains that can provide concepts of relevance to disaster response planning. The 26 ontologies were identified on the premise that crisis information systems are predicated on 11 subject themes:

- Resources: an area referring to the material and human resources that are available or potentially available for assignment to incident operations (e.g. vehicles, warehouses, tents).

- Processes: procedures and tasks that emergency management personnel need to perform in order to prepare for and respond to a crisis event (e.g. search and rescue, traffic control, evacuation).

[1] An ontology is a formal and explicit description of concepts dedicated to a particular field or domain, of properties and characteristics of those concepts and of relations between those concepts.

- People: emergency management personnel which include first responders, policy makers and disaster relief operators.

- Organizations: a variety of working groups that are in charge of crisis planning and response. It includes but is not limited to governments, military, charities and non-governmental organizations.

- Damage: an area referring to the effects of a crisis event on human, physical, natural and economic entities. Examples include injuries, missing people, missing pets, gutted buildings and burned forest.

- Disasters: a subject area concerning the classification of disastrous events or crises. Example categories include natural disasters and technological disasters.

- Infrastructure: a subject area referring to the critical physical and organizational infrastructure upon which humans heavily depend on in their normal daily life.

- Geography: a subject area providing geospatial information about disasters such as the position of a place where an incident occurs, the geo-location of displaced people and where shelters are located.

- Hydrology: information about water including the type of water body, its location, current, and how it moves across the landscape.

- Meteorology: information about weather and climate including humidity, precipitation, pressure, visibility, wind, storm and lightning, and average changes of weather during a long period of time.

- Topography: a subject area dealing with the surface shape and features of the Earth. It provides information on the terrain, ecological regime, division, and built-up area of the location where a crisis.

The subject areas identified showed two groups of concepts involved in crisis information systems: common concepts (people, organizations, resources, disasters, geography, processes, infrastructure, damage) and unusual concepts (topography, hydrology and meteorology). A summary of existing ontologies relative to the eleven subject areas identified is presented in Table 2.

Subject Area	Ontology Name	Representation Language	Downloadable	Documentation
Resources	SOKNOS	OWL-DL	No	Minimal (academic nature)
	MOAC	RDF	Yes	Online Specification
	SIADEX	Not Known	No	Minimal (academic nature)
Processes	ISyCri	OWL-DL	No	Minimal (Private wiki and in French)
	WB_OS	XML	Available on Request	Academic Nature
People	FOAF	RDF	Yes	Online Specification
	BIO	RDF	Yes	Online Specification
Organizations	ERO2M	N/A	No	Academic Nature
	IntelLEO	RDF	Yes	Online Specification
	Organization Ontology	RDF	Yes	Online Specification
Damage	HXL	RDF	Yes	Online Specification
Disasters	EM-DAT	N/A	Online Query	Classification of disasters available
	UNEP-DTIE	N/A	Online Query	Classification of disasters available
	Canadian Disaster DB	N/A	Online Query	Classification of disasters available
	Australian Gov ATG Disasters DB	N/A	Online Query	Classification of disasters available
Infrastructure	PSCAD	N/A	No	Minimal (Academic Nature)
	EPANET	N/A	No	Minimal (Academic Nature)

	OTN	OWL	Yes	Specification Available
Geography	GeoNames	RDF	Yes	Online Documentation
Hydrology	Ordanance Survey Hydrology Ontology	OWL	Yes	Online Documentation
Meteorology	NNEW Weather Ontology	OWL	Yes	Online Documentation
Topography	USGS CEGIS	OWL	Yes	Online Documentation
	Ordanance Survey Buildings and Places Ontology	OWL	Yes	Online Documentation
	E-Response Building Pathology Ontology	OWL	Yes	Not Available
	E-response Building Internal Layout Ontology	OWL	Yes	Not Available
Other	AktiveSA (multi-domain)	OWL	Yes	Not Available

Table 2: Existing Ontologies Conducive to Disaster Management

Of the ontologies designed originally for crisis management, very few are formally represented and publicly accessible. Crisis oriented ontologies such as SOKNOS (Babitski et al, 2009), ISyCri (Benaben et al, 2008) and WB-OS (Chen-Huei et al, 2011) are formally represented but are not publicly available. Moreover, no formal ontologies are specifically intended for representing the emergency and disaster domain. The existing ontologies that represent the disaster subject area are mainly in the form of database schemas, which are not (presently at least) publicly available.

It is noteworthy that most of the ontologies identified (20 out of 26) describe concepts in a single subject area. However, a few ontologies represent multiple subject areas

(QI-3). These include SOKNOS (three areas: resources, damage and disasters), MOAC (four areas: resources, processes, damage and disasters), SIADEX (de la Asunción, Castillo, Fdez-Olivares, García-Pérez, González and Palao, 2005) (three areas: resources, processes and geography), ISyCri (three areas: processes, damages and disasters), HXL (four areas: damage, geography, organization and disasters), and AktiveSA (seven areas: geography, infrastructure (transportation), meteorology, processes, resources, organizations and people

This section provides an overall picture of the subject areas and how ontologies are presently represented and used in crisis management systems. Ontologies do offer superior capabilities in querying information (as opposed to a standard SQL based application) nonetheless it is apparent from this scoping exercise that 65% of the existing ontologies are semantically interoperable. In the main differences between ontologies culminating in their semantic interoperability can be at language level (mismatches in syntax or semantic) or at ontology level (similar terms describe different concepts, different terms describe the same concepts, different level of granularity, or different coverage of domain). In order to exchange information and knowledge between diverse ontologies, the COBACORE data framework will need to be able to reconcile them. The most common approaches to reconciliation are ontology merging and ontology alignment (Hameed *et al.*, 2004). In merging, a common ontology is created by unifying all information from all source ontologies. On the other hand, in aligning, the source ontologies are made consistent and coherent with one another but kept separately. Interoperability is enabled by a set of alignment statements which specify the relations between the ontologies and is predicated on semantic intelligence modelling (Section 2.3).

2.3 Shared Vocabularies and Commonality of Language Terms

The COBACORE platform aims to improve communication between affected community, responding community and professional responders. Communication is the most commonly cited issue in multi-agency response, not only in terms of the interoperability or disruption of communication systems, but also in terms information sharing and creation of interpersonal understanding. This interpersonal understanding requires agreement on the meaning of the terms that are used. There is however no general consensus on definition of basic concepts such as disasters, hazards, risks, vulnerability and so on. Professionals from different disciplines use alternative meaning of these concepts and in turn different methods of measuring them. Although the academic community, in particular, is still debating the details of such distinctions emergency responders would in the main argue that such Disagreements lie in minor aspects of definition and pose few operational difficulties. Nevertheless there are situations where the lack of agreement on a concept creates problems in practice. An example of such problems is reporting and archiving events. For instance, to include an event in the International Disaster Database it should, at least, fulfill one of the following criteria: (1) 10 or more people reported killed, (2) 100 people reported

affected, (3) declaration of a state of emergency, and (4) call for international assistance. However different cultures perceive and therefore classify extreme events with differing degrees. A similar problem is apparent when it comes to the classification of disasters such as natural, human-caused, technological and forth across different jurisdictions.

Gaining consensus in terms of the ‘human’ usage and understanding of terminology associated with disaster incidents does not necessarily however facilitate interoperability from a technical perspective within the confines of disaster management response, different domain vocabularies are used by different crisis information systems. The autonomous nature of crisis information systems presents a challenge to exchanging information efficiently since the semantics of the data can be heterogeneous and not easily assimilated. For example, the word ‘Person’ can have different meanings - a ‘displaced person’, ‘recipient of aid’, or ‘victim’. The relative absence of a ‘common language’ or ‘dictionary of agreed terms’ across jurisdictions and amongst different entities within the professional responder community has often culminated in heterogeneous terms and concepts which are not interlinked or lack comprehensive consistency which in turns militates against “semantic interoperability” across disaster recovery systems. The term ‘semantic interoperability’ refers to the ability of computerized systems to communicate data with a unified meaning.

Within the confines of disaster response and emergency management there has been marked progress in recent years pertaining to the development of shared vocabularies. This has been driven due to the limitations and ineffectiveness of information systems to communicate effectively during recent crises such as Hurricane Katrina and the Tsunami in Haiti.

The Management of a Crisis Vocabulary (MOAC) for example is a lightweight vocabulary aiming to provide terms to enable practitioners to relate different "things" in crisis management activities together as Linked Data. The initial MOAC terms originated from the Inter Agency Standing Committee (IASC), Emergency Shelter Cluster in Haiti, UNOCHA 3W Who What Where Contact Database and Ushahidi platform. Efforts have been made to involve a number of international humanitarian and crowdsourcing volunteering communities to authenticate MOAC usability, functionality and structures. Since, the vocabulary is still evolving, necessary changes based on expert recommendations will be carefully documented and necessary changes will be continuously updated in the same platform.

HXL is another shared vocabulary for the humanitarian domain and is machine queryable. It redefines the classes and properties shared with MOAC, instead of importing them, to make sure that the vocabulary is under control of UN OCHA. It also reuses classes and properties from broadly used vocabularies such as FOAF, the Open Geospatial Consortium's GeoSPARQL ontology, DC, and SKOS. HXL was re-developed using a Linked Data approach for a number of reasons:

- **Extensibility**. HXL can only provide the core classes and properties for the domain. Publishing it as an RDF vocabulary allows data publishers to extend it according to their needs.

- **Reuse of external data sources**. Many sources in the Linked Data cloud contain information that is also useful in a humanitarian context, such as geographic or demographic data.

- **Semantic annotations**. An XML schema strongly focuses on syntactic interoperability. In Linked Data, RDF provides the syntax, and shared vocabularies define the semantics of the shared data.

- **Standardized API**. Data access works through standard HTTP requests, and the data can be queried in the SPARQL query language.

- **Inference capabilities**. The structure of the shared vocabularies supports inference (also referred to as reasoning) on the data.

MOAC and HXL encapsulate many of the visions and ideologies conceptualised by the OASIS[2] Emergency Management (EM) Technical Committee (TC). The EM-TC creates vendor-neutral and platform agnostic standards for organizations and agencies to more easily exchange emergency information. The EM-TC welcomes participation from members of the emergency management and response community, developers and implementers, and members of the public concerned with disaster management and response. The Oasis EM-TC have been responsible for the development of Emergency Data Exchange Language (EDXL) a broad initiative which is designed to create an integrated framework for a wide range of emergency data exchange standards to support operations, logistics, planning and finance. In the development of EDXL, a stake-holder driven approach ensured wide acceptance of term definition between practitioners, government representatives as well as the software industry (Subik et al, 2010). The efforts of the EM-TC to develop a commonality of language for Emergency Management globally continues to be a work-in-progress. Sustained progress has nonetheless been made by the EM-TC internationally to improve the interoperability of information systems and to develop commonality of language. Pertinently, within the confines of the COBACORE data framework development, the Emergency Management TC approved Emergency Data Exchange Language Situation Reporting (EDXL-SitRep) in April 2013. In order to ensure, continuity of interpretation and cross-

[2] OASIS is the Organization for the Advancement of Structured Information Standards, a not-for-profit, international consortium that drives the development, convergence and adoption of open standards for the global information society. OASIS promotes industry consensus and produces worldwide standards for security, Cloud computing, SOA, Web services, the Smart Grid, content technologies, emergency management, eGovernment, and many other areas

jurisdictional application it is important that the standards and Emergency Management language terms being advocated by the EC-TC are integrated into the final data framework design. Key standards developed thus far by the EM-TC include: EDXL Common Alerting Protocol (EDXL-CAP).

- EDXL Distribution Element (EDXL-DE)
- EDXL Hospital AVailability Exchange (EDXL-HAVE)
- EDXL Resource Messaging (EDXL-RM)
- EDXL Reference Information Model (EDXL-RIM)
- EDXL Situation Reporting (EDXL-SitRep)
- EDXL Tracking Emergency Patients (EDXL-TEP)

These standards will serve as an invaluable starting base for data framework development. However, the functions and features required from the COBACORE platform as detailed in Chapter three of this document as well as the need to capture, interoperate and analyse inputs from the affected community as well as professional responders will necessitate the application semantic intelligence modelling in order to ensure language commonality and automatable interpretation.

2.4 Semantic Intelligence within Disaster Recovery

The application of Open Geospatial Consortium (OGC) and Web Feature Service(WFS) technologies has markedly improved the sharing and synchronization of feature-level spatial and non-spatial information within the confines of disaster recovery and reconstruction. Nonetheless, existing implementations of OGC WFSs only serve to emphasise syntactic data interoperability via standard interfaces, and cannot resolve semantic heterogeneity problems encompassing both spatial and non-spatially referenced data – as will be necessary within the COBACORE platform. As such it is problematic to perform an intelligent content based search and users cannot correctly utilize the discovered WFSs without additional human assistance or programming. To overcome the aforementioned problems the COBACORE data framework will utilise ***Semantic Intelligence Modelling.***

Semantics is the study of meaning. It focuses on understanding the relation between signifiers, like words, phrases, signs, and symbols, and what they stand for, their denotation. This problem of understanding has been the subject of many formal enquiries, over a long period of time, most notably in the field of formal semantics. The formal study of semantics intersects with many other fields of inquiry including linguistics, lexicology, syntax, pragmatics, etymology, as well as philology, communication, and semiotics; and in recent times, software engineering.

In linguistics, semantics is the study of interpretation of signs (or symbols) used in agents within particular contexts. Therefore, sounds, facial expressions, body

language, and proxemics have semantic (meaningful) content. In written language, things like paragraph structure and punctuation bear semantic content.

In software engineering semantics is about the meaning of data in terms of the interrelationship in data, resulting in semantic data model. A semantic data model is an abstraction which defines how data items (i.e. entities or symbols) relate to the real world. Semantic data models are characterised by the use of directed graphs in which the vertices denote concepts or entities in the world, and the arcs denote relationships between them. Semantic data modelling techniques include Resource Description Framework (RDF) and Web Ontology Language (OWL).

The multi-faceted and inter-disciplinary nature of crisis response is wholly conducive to semantic intelligence modelling both in terms of developing relationships and linkages between datasets relative to the disaster incident, to enhance usability in terms of the human-technical interface in order to develop commonality of understanding.
Semantic Data Modelling will be used for the following purposes in COBACORE:

- Planning of Data Resources: A preliminary semantic data model will be used to provide an overall view of the data required to provide situational awareness in the confines of a disaster incident. The model can then be analysed to identify and scope projects to build shared data resources across the professional responder community.

- Building of Shareable Databases: A fully developed semantic data model can be used to define an application independent view of data which can be validated by users and then transformed into a physical database design for any of the various DBMS technologies. In addition to generating databases which are consistent and shareable, development costs can be drastically reduced through semantic data modelling.

- Integration of Existing Databases: By defining the contents of existing databases with semantic data models, an integrated data definition can be derived. With the proper technology, the resulting conceptual data model can be used to control data processing in a distributed database environment.

The sematic intelligence specifications are explored in-depth in section 4.0. This serves to identify the role of semantics in terms of search functionality and analytical capacity and the features and functions predicated on the 'No Regrets' Feature list emanating from WP3.

3.0 COBACORE DATA FRAMEWORK: SPECIFICATIONS AND USER REQUIREMENTS

The disaster response community is moving towards efforts to define and more effectively exploit ICT utilisation during the response and recovery phases of major incidents/emergencies. The focus has moved from just improving voice-data-network-level technologies for communication into harnessing new information-level technologies to cover all phases of crisis management. This includes information infrastructure for incident message routing and standard languages for conveying the semantics of emergency warnings and resource and task management.

Many of the existing data frameworks centre around the system-of-systems ideology, seeking to integrate the often autonomous and heterogeneous information systems utilised across the professional responder community. The ISyCri framework in France for example adopts a UML meta model of crisis situation management and its corresponding OWL ontology in order to federate and derive meaning and facilitate interoperability across the existing systems of different organisations (police, medical units, NGOs etc). The project advocates the use of a mediator - Mediation Information System (MIS) between partners in order to support interoperability between heterogeneous systems (Benaben et al, 2008).

Paradoxically, whilst extensive research has been undertaken regarding the sharing of information between professional responders, less attention has been paid to the sharing and coordinating of information flows between the affected community and the professional responder community.

Advances in ICT including the rapid expansion in the use of social media has opened up new avenues pertaining to data collation as well as expediting communication flows and the extent of outreach. The consensus amongst the professional responder community is that the use of social media and associated technologies within the confines of a disaster incident has not yet been fully exploited with issues pertaining to the most effective means of capture, co-ordination, utilisation and validation of this increasingly important communication medium in order to bolster situational awareness within crisis information systems.

To date, social media has been predominantly utilised as a means of information dissemination (cascaded down from professional responder community/civil contingency co-ordinators) and has failed to be fully exploited as a credible and contemporaneous means of data capture. This in part has been attributable to the inability of existing disaster information systems to effectively capture social media and translate what can be highly voluminous unstructured data sets into meaningful validated information. The COBACORE platform is intended to expand the focus of information exchange to encompass the flow of data and information between the professional responder community, the affected community and volunteer responders.

In addition to the deficiencies which are considered to pertain to the most effective exploitation of social media there is growing recognition that more needs to be done to effectively harness the capacities of local communities within the confines of the relief effort. Thus far the capacities of communities have not been fully utilised. There is ineffectual co-ordination of volunteer responders, presently many data systems including social media sites formed in the aftermath of a disaster are sporadic in nature and do not always serve the best interests of the relief effort. Additionally, there is real capacity for duplication and repetition of effort (across multiple sites) which detracts from the efficiency and effectiveness of those who are seeking to help. A key objective of COABACORE is to more effectively harness the skills and capacities of volunteer responders and to ensure more effective alignment of capacities of this important grouping with the needs of the affected community. Facilitating this will necessitate more effective communication flows between the aforementioned stakeholder groupings (Figure 2).

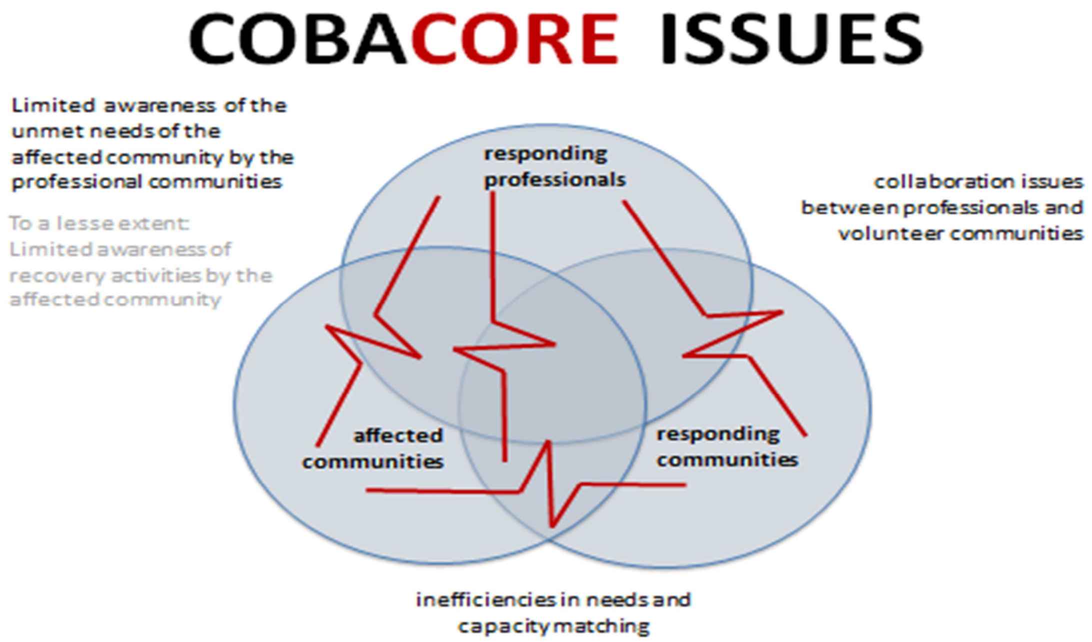


Figure 2: Key Data Needs Across Stakeholder Groupings

4.0 FUNCTIONAL REQUIREMENTS OF COBACORE DATA FRAMEWORK

Data frameworks must be designed to meet end-user needs and expectations, ranging from data acquisition through data entry, validation, reporting, and long-term analysis. Data modelling is the methodology that identifies the path to ensure conformity with user requirements. The focus should be to keep the overall data framework structure as simple as possible while still adequately addressing project goals and objectives. The following broad questions are a good starting point in devising a roadmap to guide framework design:

- What are the data frameworks objectives?
- How will the data framework assist in meeting those objectives?
- Who will use the data framework and what tasks do those individuals need the database to accomplish?
- What information will the data framework hold?
- What is the granularity of information the data framework will hold and what are the format and features of the datasets?
- Will the data framework need to interact with other data information systems and applications? What accommodations will be needed to facilitate this?

In order to answer these questions, we need to reflect on the feature development process of the COBACORE platform. However, because of the variety in applicable use cases for our platform, and the overall novelty of the approach, we are not in a position to settle on the final feature set of the platform early on in the project. The original proposal gave a general direction for the platform functionality, but fell short of a functional specification of the eventual platform. Through domain analysis and stakeholder evaluation sessions, we need to uncover valuable platform functionalities, and we need a project development strategy that can accommodate for this step-by-step process.

We have adopted a simple incremental platform development process. The process derived major platform functions from identified domain issues. From asserted functions, features are derived in a number of iterations, starting with the definition of a core feature set. Features should not be regarded as a fine-grained functional specifications, but rather as blueprints that need to be satisfied by the eventual implementations. The features give direction to the implementation by proposing logical structures, interaction patterns and suitable interface elements. The actual design and implementation choices result from considerations by the various COBACORE project teams, from their own perspective.

In subsequent phases, new features can be defined that build upon previously implemented features, and thus add new capabilities to the platform. Not all proposed features might be realisable during the project due to time constraints or technical

limitations, so certain features might need to be realised in follow-up projects. Figure 3 illustrates this feature development process.

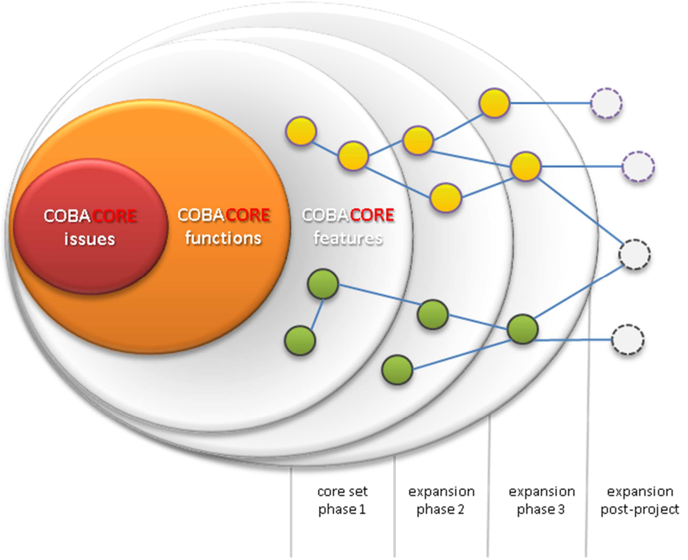


Figure 3: The COBACORE feature development process

The domain analysis, as carried out by the WP1 team, has established the core domain issues that are abound in the disaster recovery domain, and affect recovery performance. Table 3 summarises the findings in three core issues:

Issue	Description
I1	Information provision issues between the professional community and the affected community
I2	Collaboration issues between the professional community and the responding community
I3	Inefficiencies in needs and capacity matching between the affected and responding community

Table 3: COBACORE Core Issues

From these three core issues, three core functions were derived. These functions assert the major functions that the COBACORE platform should provide. 4 lists the three main COBACORE functions.

Function	Description
U1	Enhance information exchange between the professional community and the affected community
U2	Facilitate collaboration between the professional community and the responding community

U3	Improve needs and capacity matching between the affected and responding com

Table 4: COBACORE Core Functions

The three core functions do not state how the platform should behave or look, but denote the value of the platform for the target user groups. The functions thus guide the development of platform features and interfaces.

As mentioned in the introduction, the first phase of the feature development process consists of establishing a core set of features. This set has also been referred to as the 'No-Regrets list', hinting at the critical quality of these features. These core features represent the minimum functional requirements that the platform implementation should fulfil. Table 5 lists the core feature set.

id	feature name	prime user group	feature purpose
F1/AR	Actor registration	Affected community, Responding community, Professional community	Provides means for users to register actors with the platform. An actor is the bearer of a need, a capacity or an activity.
F2/NR	Needs registration	Affected community	Provides means for users to register needs related to their recovery from a disaster. They may register needs for themselves, or as representative of a group.
F3/CR	Capacity registration	Responding community, Professional community	Provides means for actors to register capacities that they are willing to provide towards the disaster recovery.
F4/NCO	Needs & capacity overview	Affected community, Responding community, Professional community	This feature provides means for users to view registered needs and capacities, and perform various filtering and sorting operations to establish an informative overview.

F5/NCM	Needs & capacity matching	Affected community, Responding community	Provides means for users to establish matches between registered need profiles and registered capacity profiles. After a match has been determined, it can be effectuated as an activity.
F6/AAO	Actors and activities overview	Affected community, Responding community, Professional community	Provides means for users to view registered actors and activities, and perform various filtering and sorting operations to establish an informative projection.
F7/BSO	Baseline situation overview	Professional community	Provides an overview of the situation in the affected area through projection of baseline data and information.
F8/BRV	Basic recovery views	Professional community	Provides the user with options to build data views that are relevant to understanding the recovery process, and thus provide a basis for further analysis and plan-definition.
F9/BIE	Basic information exchange	Affected community, Responding community, Professional community	Provides means for users to exchange information (documents, data, digital media), centered around mutual capacity profiles and activities.
F10/ACT	Activity registration	Responding community, Professional community	Provides means for users to register an activity with the platform. An activity is a recovery effort in which actors partake.

Table 5: The COBACORE Core Feature Set

A more extensive description of the core feature set can be found in D3.1, and on the project Sharepoint site. In order to conform with the 'No Regrets Features' list a number of key assumptions need to be considered within the confines of the data framework design:

Pre-Crisis Implementation

There is an expectation that the research framework being proposed within COBACORE will be in position prior to the onset of a disaster event. It is considered impossible to implement and set up a data model within the dynamic nature of the disaster. Much of the baseline information profiling an area is 'static' in nature and can be accessed/assembled pre-crisis. This affords the benefits of instant access for

emergency responders in the event of a crisis. Moreover, a number of research studies (Bharosa et al, 2010; Turoff et al, 2004; Venkatesh et al, 2003) have demonstrated that Inter-Organizational Information Systems (IOIS) that are not used on a regular basis prior to a disaster will never be fully utilized or bring significant operational benefits in an actual disaster. Pre-crisis awareness, application and usage of IOISS systems is essential in order to promote system familiarity and confidence from an end-user perspective be they professional responders or members of the affected community. It is important that the COBACORE platform has sufficient functionality and regular updating to warrant usage in the non-crisis situation (See SF72 framework).

Framework Interoperability

When considering the design features of the COBACORE data framework a key requisite was the recognition that the system will not exist in isolation. Although a number of stand-alone systems do exist including web-based and client/server solutions, such as WebEOC (ESi, 2007) and L-3 CRISIS (Ship Analytics, 2007) the requirement for all parties involved in crisis management to use a single (and often centralised) system has ultimately hindered uptake. The COBACORE data framework must recognise and embrace interoperability through recognition of and conformance with common standards, both between different implementations of Crisis Information Management System (CIMS), as well as between CIMS and other types of software used by the professional responder community. However, whilst many countries have implemented standardised terminology, principles and command structures for crisis management by developing their own Information Management Systems (IMS) (such as AIIMS in Australia) there has been limited adoption of standard formats for information sharing between information systems such as CIMS. It is important therefore that the COBACORE data framework is designed to facilitate interface compatibility with existing crisis management information systems in order to compliment existing crisis information systems as a prerequisite for multi--jurisdictional transferability and cross-disciplinary application.

Scalability and Transferability

The architecture is designed to support different deployment models at different scales to facilitate gradual growth and expansion as needed by responders. Availability of relevant and timely information is of utmost necessity in conducting and handling of large amounts of information. Nonetheless it is apparent that many of the datasets compiled within the confines of the baseline position have currency respective of the nature of the disaster/emergency. Therefore consideration needs to be given to the 'tagging' or classification of incidents to ensure that the most relevant datasets are elevated. This will safeguard against information 'overload' to prospective COBACORE end-users.

Operational Resilience

Because CIMS systems come under heaviest load when a disaster occurs and the immediate response phase, they have to operate in challenging external conditions. Network connections might be intermittent, network nodes have to be able to join and disconnect at will, and information has to be accessible to end user terminals with limited resources, such as PDAs and mobile phones. This presents challenges to the data framework design and the operational features compatibility.

Data Ownership and Maintenance

The COBACORE data framework has been populated with a series of base layer datasets much of which are open source and current. This premise is to demonstrate the 'proof of concept' and the prospective added value of COBACORE. However, it must be acknowledged that a framework of this nature in order to bring added value to the disaster recovery process will require ownership in terms of regular updating of base layer data and maintenance to ensure conformity with advances in technology.

Once the features and functions of the COBACORE model have been agreed the next stage in the process from a WP2 perspective is to assemble the necessary datasets and devise a data framework conducive to the operational features and functions required.

5.0 THE COBACORE DATA FRAMEWORK

The primary objective of D2.2 is to detail the technical considerations and specifications as well as the models and hierarchies of the COBACORE data framework in order to facilitate compliances with the key functions and feature specifications developed in WP3, to support the management, visualisation, mining and model development undertaken in WP4 and in conformance with the needs and expectations of the evaluation exercises designed and developed in WP5. It is intended that the COBACORE data framework will utilise as far as is practically possible the existing 'state of the art' methodologies and ontologies presently adopted within the confines of disaster response, recovery and reconstruction identified in section 2. It should nonetheless be acknowledged at this juncture that the focus and scope of the COBACAORE platform 'sits outside' the features and functions of many existing crisis management information system. As such it is envisaged that whilst existing ontology's provide a solid base for initial framework development the 'bespoke' nature of COBACORE will necessitate an expansion of these existing ontologies as well as the utilisation of semantic intelligence modelling to interlink them. The identification and reconsolidation of existing ontologies relative to COABCORE as well as the sematic architecture of the COBACORE framework is detail in sections 4.3 and 4.4).

5.1 Methodology

Data modelling is an iterative process; there is no one 'correct' way to model a domain. However, the goal of this initial iteration is to produce a data framework structure that is flexible enough to accommodate change, both, as we progress through the project, and, as the requirements of different scenarios change over time. A technical advisory board has been established to inform and guide the development of the framework specifications and this will be complimented through a series of on-going discussions with prospective end-users encompassing professional responders and the affected community in order to refine and validate key data requirements and specifications.

There are two main types of approaches when attempting to derive a conceptual data model from a domain: a bottom up and top down approach; we will be taking a 'hybrid' approach, aiming to combine the best of both approaches (Xu, 2007).

Using a hybrid approach we will able to iteratively decompose the top level concepts while identifying missing vocabularies in the COBACORE domain, and, at the same time, making sure that our bottom level data sources can be subsumed by the framework. In this way, the model will be sufficiently abstract and flexible to deal with changing environments but, at the same time, validating the model by ensuring all the datasets deemed critical to inform the crisis response (premised on feed-back from elevation exercise and engagement with prospective end-user) are catered for.

The data modelling exercise will utilise, where appropriate, existing ontologies and vocabularies. This 'reuse' approach will maximise the interoperability of the data

framework, provided that the ontologies used are 'fit for purpose' i.e. stable and prevalent. The 'state of the art' ontologies identified in 'Ontologies for Crisis Management: A Review of State of the Art in Ontology Design and Usability' (Liu, 2013), serve as a good base for creating a 'fit for purpose', 'state of the art', data framework.

Building upon the work of (Liu, 2013) we can describe how the ontologies, identified in their research as commonly used concepts, can be applied to the COBACORE concepts, what deficiencies exist, and how these deficiencies can be addressed.

As identified in section 2 and (Liu, 2013), there exists several 'state of the art', but 'incomplete', ontologies for describing a community based disaster and recovery domain. Where incomplete vocabularies or ontologies are found, the decision is made to either extend the existing ontology/vocabulary or to attempt to create a new, complete, entity using the 4W (who, what, where, when) approach advocated for modelling the humanitarian domain (Limbu, 2012).

The COBACORE data framework will change, and have the ability to change, over several iterations relative to required features and functions and in light of findings derived from both the intermediate evaluation exercise and the final evaluation exercise being designed in WP5.

Ultimately the 'success' of the data framework will be determined by:

- Its ability to embrace the functional requirements (outlined in section 3).
- Its interoperability with existing crisis management information systems by being standards driven and semantically linked.
- Its ability to be 'state of the art' by being semantically interlinked with a wide range of existing data sources deemed relevant to crisis response management by prospective end-users.

Its ability to be used as the building block for WP3, WP4 and WP5. In order for WP3, WP4 and WP5 to move forward they require the data framework to be in a format which is usable for the concept and platform development activities that are to be delivered. The main output from this section will be a conceptual data model which will adequately facilitate the creation of use-case diagrams in line with function and feature requirements (for WP3) and a physical data model which is fully interface compatible with the COBACORE End-User platform (WP4). Furthermore the framework must be developed to an operational standard conducive with the intermediate and final evaluation exercises (WP5) as well as being suitably robust and 'useable' to attain credibility and viability as a 'proof of concept'.

5.2 Data accessibility

An overarching goal of COBACORE is to enable a flexible loosely-coupled platform linked by industry-standard interfaces, and open interchange of meaningful data and metadata with ease of use. Therefore it is essential to the platform success to support establish and spreading open protocols for accessing and distributing heterogeneous data sources.

When dealing with access protocols we can distinguish in general between

i) multi-domain data sources which may be spatially referenced directly or indirectly on a coarse scale, e.g. at district, city or region level ; and
ii) spatial datasets, i.e. map sources, with explicit geo-referencing using well known coordinate/projection systems.

For the first type of data sources that we will refer to in general as *databases*, we can differentiate in turns between structured and unstructured information. The second one is meant for information which does not have a defined structure, but generally comes packaged in objects such as files or documents. Typical access protocols used to be considered for this kind of data sources are content access/delivery protocols such as HTTP or FTP which are commonly present today in any electronic services platforms. Also, high performance, peer-to-peer access protocols such as BitTorrent will be considered when dealing with community-shared, large-size information or content sources.

Regarding access to structured databases, COBACORE will first support well established protocols present in most DBMS (Database management Systems) and related to the SQL (Structured Query Language) technology, such as ODBC (Open Database Connectivity) or JDBC (Java Database Connectivity). When needed, specific protocols for accessing given structured databases will be supported, such as LDAP (Lightweight Directory Access Protocol) for accessing organisations information.

Being an essential aspect of the COBACORE approach to data, the platform will support existing and upcoming access protocols suitable to access semantic databases. In general terms this means enabling support for accessing RDF objects and the SPARQL query language. In addition, there are some upcoming industry standards in this field such as OData (Open Data Protocol), promoted by Microsoft and based on the EDM model (Entity Data Model) rather than RDF; or the LDP (Linked Data Platform) protocol promoted by W3C , based on RDF). Other vender-specific widely spread formats such as Google Fusion Tables will also be considered.

Regarding access to spatial data sources, the main frontend protocols to be supported will be the widely used OGC WMS (Web Map Service) and OGC WFS (Web Feature Service). In order to access raster data in general we consider it is sufficient in general to use a content access protocol such as HTTP or FTP, although some specific solutions could be also added for some particular sources, e.g. the OPeNDAP (Open-source Project for a Network Data Access Protocol) protocol for accessing weather or other earth-sciences related data.

The volume of data means that, where possible, the data should be retrieved 'from source' and not retrieved and stored locally in a private repository. Some information sources will require local storage due to security, availability and performance constraints but, in order to maximise the integrity, flexibility and maintainability of the system, off-site data should be used where possible.

To achieve the goal of using off-site data, the data framework has been designed in such a way as to allow access to these disparate, off-site and on-site, data sources in a generic manner using 'state of the art' web-based semantic technologies and ontologies.

LOCAL DATA

For performance, accessibility and security reasons some data will need to be retrieved and hosted locally. The types of data that will need to be hosted locally will include map layer types, spatial/non-spatial data stores, and metadata definitions.

Map layer sources can be hosted locally using a geospatial server such as GeoServer (http://geoserver.org/), MapInfo or ArcGis.

Other data sources will take the form of either spatial or non-spatial datasets. Spatially related datasets can be hosted and queried using a spatially aware relational database management system (RDBMS). Open source options such as PostGIS (http://postgis.net/) exist which allow spatial data to be stored and queried.

Non-spatial data can be stored locally using a structured or non-structured data technology such as an RDBMS or NoSql (Not only sql). The data from these sources can then be exposed using a common query language such as SPARQL or a Linked data REST API.

REMOTE DATA

Certain sources of baseline data that do not need to be imported, stored and maintained locally e.g. public government data available through a public API or triple store, can be 'plugged into' COBACORE queried and accessed ad-hoc. These types of sources can be utilized in the platform through a number of channels:

Public WMS, WFS servers, KML, GEORSS, GWC. For this type of source only the location of the connection details and metadata for the layers are required. E.g. the UK government, through INSPIRE, have made a public WMS server available where layers can be dynamically accessed.

Public API. Some public data e.g. open.gov.uk can be accessed by means of a public api. The data from these sources can only be used by creating 'adapters' to allow access and expose of this data in a format that the components of COBACORE will require. This data can then be exposed to the COBACORE platform through the means of a common API/Query language.

SPARQL/GEOSPARQL. However some of the public api data sources also expose their data using semantic web technologies that are query able using languages such as SPARQL/GEOSPARQL e.g. data.gov.uk, DBPedia, UNOCHA FOD/COD, world bank indicators. If performance or availability becomes an issue with any of the endpoints

the RDF data dump should be hosted and exposed locally using self-provisioned SPARQL endpoints or a REST Based API.

Public data stores. COBACORE will also have access to publically and non-publically accessible data stores e.g. remote RDBMS or Google Fusion tables. This data should ideally remain remote, but in order to expose this data to the rest of the COBACORE platform the metadata needs to be described and stored before it can be used.

5.3 Data Security

The COBACORE platform will include as core functions those needed to handle the lifecycle of the different types of data sources supported. This include, as a self-standing function, the capability to adequately protect information. Every piece of sensitive data will have corresponding and indivisible security metadata at the instant it is created, which declares rights holders and protection requirements for the different potential usages across and along the system until it is eventually eliminated. Different security procedures will be explored and selected, both for databases and spatial data sources, to enforce separation of duties and data protection enabling a high-integrity security system able to properly prevent leakage or undue modification of sensitive information. Logging of security events on data (access, modification, transmission) will be also specified in order to verify that declared security requirements are respected.

The data management subsystem in the platform will include different data repositories (for structured, non-structured and map based information) which are handled by corresponding drivers. One central data manager module, connected to the different repository drivers, will offer uniform access to end users to the different information sources hosted. We can distinguish there different segments in the path from repositories to user terminals where data protection is needed. The first two parts refer to the driver access to the physical repository and to the connection between the data manager and the repository drivers. The protection solutions to be used will depend on specific tools selected for the implementation of the data subsystem, and in particular will allow a fine-grained selection of protection levels (e.g. hashing or encryption) depending on individual data objects and collections and offer both data confidentiality and integrity services.

The last part of the data path refers to the connection between the end users and the COBACORE platform, this is, to the protection of information flows provided by the platform. The security approach here will consist on properly authenticating end users and on a fine-grained handling of access rights for the different data collections and objects; which is compatible and possible to integrate with the protection techniques used within the platform. The solutions to be selected will be based on industry standards such as WebID, OpenAuth, PKI, TLS or IPsec and on the use of trusted existing services, They will also observe the peculiarities of the COBACORE scenario, such as the need to consider particular mechanisms for large-scale voluntary community contributions; or to handle security aspects of automated processing

chains, e.g. how rights might be created or change upon data processing. As an example, for the case of spontaneous voluntary community contributions, a random key distribution procedure is devised so that privacy and confidentiality rights are properly preserved. Also, specific solutions will be deployed for semantic information such as WAC (Web Access Control vocabulary), which handles access rights for RDF sources.

Information security elements will be specified at design time and further implemented across the different system components and platform levels, protecting in particular personal sensitive information against unauthorized access, disclosure, alteration or destruction.

Predicated on the 'No-Regrets' feature a prototype data framework has been conceptualised. The prototype is presently in the early development stage and will continue to evolve relative to the features and functions emanating from WP3 and conducive to the operational requirements being designed within the evaluation exercises (WP5). Utilising the existing ontologies it is possible to combine heterogeneous data sources using semantic intelligence. A series of screen shots from the prototype are presented depicting the conformance with the No Regrets Feature (Section 3.0) in order to verify the underpinning data specifications and requirements and to inform and guide semantic development in terms of defining data relationships in order to transpose valuable and meaningful information.

5.4 COBACORE Data Model Architecture

The description of the data model will begin by identifying all the core concepts related to the disaster recovery and reconstruction domain. Working from the commonly used concepts in (Liu, 2013) we can start to integrate these concepts into the COBACORE model, renaming where appropriate, and grouping the concepts into hierarchical categories.

What follows is an outline or the core concepts and concept groupings in the domain and a description of why the domain could be modelled in this way.

5.4.1. Base concepts

Two core concept groupings can be identified, within the disaster domain, to encompass all concepts of the domain. These groupings will encompass all the commonly identified concepts in the disaster and recovery domain identified in (Liu, 2013). The two groupings identified are the physical situation and the actors/community involved, directly or indirectly, in the domain i.e. the factors and actors. These concept groupings are described in more detail and then decomposed in the following sections to incorporate the more commonly used concepts.

5.4.1.1. Factors

The factors concept is useful in categorizing all the physical and non-physical elements that make up 'the situation'. Using this grouping we can now the commonly used core concepts of (Liu, 2013) to this group and also add any concepts that may be missing in order to model the COBACORE domain. Some examples of the factors include physical factors (such as meteorology, topography, hydrology and infrastructure) and the social indicators (economic indicators, aid effectiveness, demographics etc.).

The situation will change over time; 'meta' concepts such as 'time' and 'location' will be expanded upon later in this section to show how they apply across concepts.

5.4.1.2. Actors

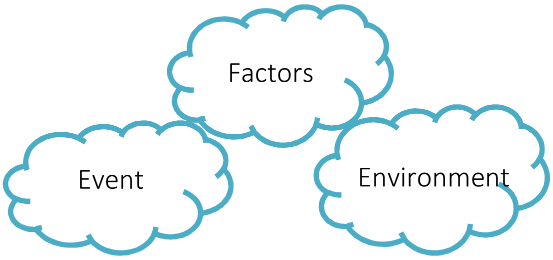
The second main concept grouping describes all the 'actors' involved in the domain.

The term 'actors' encompasses both the groups, and individuals, involved, directly or indirectly, in the domain. Both groups and individuals can be said to have needs, capacities and activities.

Every individual is a member of an 'actor' grouping as they will have a characteristic that places them in one or more group. For example, for a particular scenario, it may be necessary to identify a group based on the age (over 60s) and location (further than 3 miles away from the nearest hospital). The flexibility of the data model allows us to dynamically create groups from a collection of individuals based on their characteristics. Some groups will be predictable e.g. the Red Cross group will consist of all the individuals with a characteristic of the profession organisation being 'Red Cross', however for changing scenarios, with a myriad of competing factors and actors, a flexible and dynamic data model is required.

5.4.2. 'Factors' Core Concepts

The 'factors' consist of the disaster details and the environment it is operating in e.g. what happened, and the environment in which the disaster is operating e.g. the current/projected weather, flood plains etc.



5.4.2.1. Event

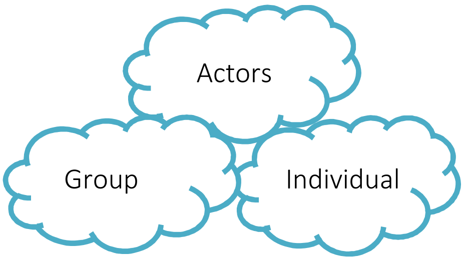
The details of the disaster i.e. the who, what, where, when (and why) of the situation.

5.4.2.2. Environment

The social and physical environment in which the situation is currently operating e.g. the infrastructure, hydrology, meteorology, and topography of the situation. The social environment is also included here and provides indicators from existing data on the social makeup of the environment relating to the situation.

5.4.3. CORE ACTOR CONCEPTS

The 'actor' concept comprises of the individuals and groups of individuals who are relevant to the domain i.e. it encompasses all of the responders, affected population, involved organizations etc. All groups and individuals have associated capacities, activities and needs.



5.4.3.1. Groups

A 'group' describes a collection of individuals (based on the individuals' common characteristics). The needs, capacities and activities of a group can be viewed as either a collation of the individual attributes of the grouping or described explicitly by the entity itself e.g. the International Red Cross may describe its own organizational/strategic needs, capacities and activities, without resorting to collating the needs, capacities and activities of its individual members. Both individuals and groups are described as 'actors' in the domain.

Each group is defined by a set of characteristics of the individuals that make up the group. Such characteristics include age, locality, profession, group membership, employer etc. For example, a 'local community' group can be defined as all those individuals living in a certain district within a city. Each group will have associated collective capacities, needs and activities which are derived from the individual members of the grouping. Some groupings will also have.

Every individual will be a member of at least one group, group membership can also change over time and location. Individuals may, or may not, be aware that they are part of a particular group e.g. an individual may be aware that they are part of the IRC group because they have the same employer as other individuals but unaware that they are part of an 'at risk' group due to their commonality in current location.

This approach, of allowing groups to be defined by a flexible list of characteristics, gives flexibility to the model. This flexibility is required to reconcile the data framework to the analytical functionalities required by COBACORE e.g. one particular functional requirement, described in section 3, is to enable 'Basic recovery views' which provides professional users with "options to build data views that are relevant to understanding the recovery process, and thus provide a basis for further analysis and plan-definition"; this type of dynamic understanding of the situation is only possible when groups can be defined 'ad-hoc' by professionals.

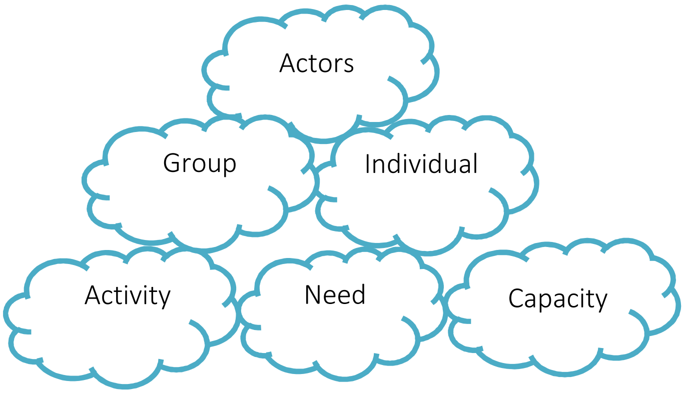
5.4.3.2. Individuals

Each individual will belong to one of more groups based on their characteristics e.g. age, gender, employer, profession, location. Each individual will have their own needs, capacities and activities which can be either taken individually or used to derive a group metric.

Every individual will be a member of at least one group, group membership can also change over time and location. Individuals may, or may not, be aware that they are part of a particular group e.g. an individual may be aware that they are part of the IRC group because they have the same employer as other individuals but unaware that they are part of an 'at risk' group due to their commonality in current location.

5.4.4. Groups and individuals' activities, needs & capacities

Both groups and individuals will have associated activities, needs and capacities.



5.4.4.1. Need

A 'need' usually expresses a necessity by a person or thing; in the COBACORE domain this definition is expanded to not only include essentials but also very important things. It is envisioned that the distinction between these 'needs' and 'very important things' will be performed by ascribing a severity attribute to the 'need' in the full decomposition exercise of the concepts in D2.3.

The COBACORE definition also only ascribes 'needs' to people and not to things i.e. all needs can be attributed to individual or groups. For example, if a bridge has been damaged, it is not the bridge itself that 'needs' repair, it is the local government group, or local community, who requires transport connectivity and thus 'needs' it repaired i.e. an inanimate object cannot have a human term such as 'need' applied to it.

The needs model maintains an overview of actual needs on different levels of abstraction and provides an outlook on how they will be met. The model also provides an overview on projected needs. It links lower order needs (local individual and community needs) to higher order strategic demands (flash appeals and strategic information frameworks that stem from humanitarian clusters).

The concept of a 'need' is not a commonly described concept in the response and recovery domain (Liu, 2013). Efforts are ongoing at UNOCHA to describe a 'need' within the HXL vocabulary (Keßler, 2009). The stated methodology for the COBACORE domain model, if a commonly used ontology cannot be found for a [need] concept, is to attempt to describe the concept using the 4W approach (who, what, where when). This 'concept' can then be expanded in D2.3 to provide implementation details such as a detailed framework for categorizing needs and the severities of that need.

Who

A need is associated with an actor (group or individual(s)).The description of a need not only needs to encompass the details of who is affected but also who initiated the need i.e. the origin of the need might not always be the affected actor.

What

The 'What' describes details about the type/categorization of need and its severity.

Where

The ‘Where’ will take account of the origin location of the need and also the affected area.

When

The ‘When’ describes details about when the need was identified and the associated timescales.

5.4.4.2. Activity

Needs and capacities may have one or more activities related to it. An activity can describe any type of action in the domain. An activity may be in direct response to a need or may be unrelated but will still have an effect on the situation. All activities will however belong to a group or individual in the actor model.

A more challenging aspect of the activity is to describe ‘why’ the activity was undertaken. The activity model should also attempt to describe ‘why’ an action was undertaken through an attempt to link an activity to a stated need. With this ‘why’ information COBACORE will be able to track the interactions and outcomes between capacities and needs and activities. However, as activities can exist without being linked to an expressed need or capacity, the activity concept will be the most be challenging in expressing ‘why’ an activity was undertaken.

5.4.4.3. Capacity

A capacity describes a potential resource that could be used to influence the situation. It not only includes the physical capacity of groups and individuals but also describes the skills, informational capacity, sphere of influence and monetary capacity of these groups and individuals.

The concept is again modelled using the 4W approach:

Who

Who has the capacity e.g. a church group, an individual (See community model).

What

What is the capacity e.g. is it certain skills, a physical resource such as a shed, information (a phone line has the capacity to provide up to date information), influence (a politician has the capacity to influence a section of the public), monetary (the local government has money available for rebuild projects.

Where

Where can the capacity be activated i.e. where is the capacity now and where can it be deployed.

When

When can the capacity be deployed and for how long.

As with needs and activities, every capacity belongs to an agent (person, group). A capacity can only be connected to a person or grouping e.g. a public grass area may be said to have a 'capacity' to hold 1000 people, but in actual fact it is the local government who owns this 'resource' and hence 'has' this capacity.

5.4.5. META CONCEPTS

Certain concepts will not be limited to a specific concept and will span both community and situation models.

5.4.5.1. Time

The changing situation and actor participations all have a temporal aspect. Past, present and future information will enable actors to take informed decisions based on all three temporal aspects.

5.4.5.2. Geo Location

Most concepts within the domain will have some sort of geospatial attribute.

5.5 Re-Use of Existing Disaster Response Ontologies

The COBACORE data framework will utilise, where appropriate, existing ontologies and vocabularies. This 'reuse' approach will maximise the interoperability of the data framework, provided that the ontologies used are 'fit for purpose' i.e. stable and prevalent. The 'state of the art' ontologies identified in 'Ontologies for Crisis Management: A Review of State of the Art in Ontology Design and Usability' (Liu, 2013), served as a good base for creating a 'fit for purpose', 'state of the art', data framework. Building upon the work of (Liu, 2013) we can describe how the ontologies, identified in this research, can be applied to the COBACORE concepts, what deficiencies exist, and how these deficiencies can be addressed. Where incomplete vocabularies or ontologies are found, the decision is made to either extend the existing ontology/vocabulary or to attempt to create a new, complete, entity using the 4W (who, what, where, when) approach advocated for modelling the humanitarian domain (Limbu, 2012). Figure 3 represents an initial reconciliation of the COBACORE concepts relative to the existing ontologies that can be used directly or extended relative to the continued refinement of the functions and features (WP3) and the necessitated requirements and operational capacities of the COBACORE platform (as determined within WP4 and WP5). The expansion and integration of these existing

ontologies relative to the operational requirements and functional specifications will form the basis of future work within the confines of the attainment of D2.3 which is due to be completed in Month 24.

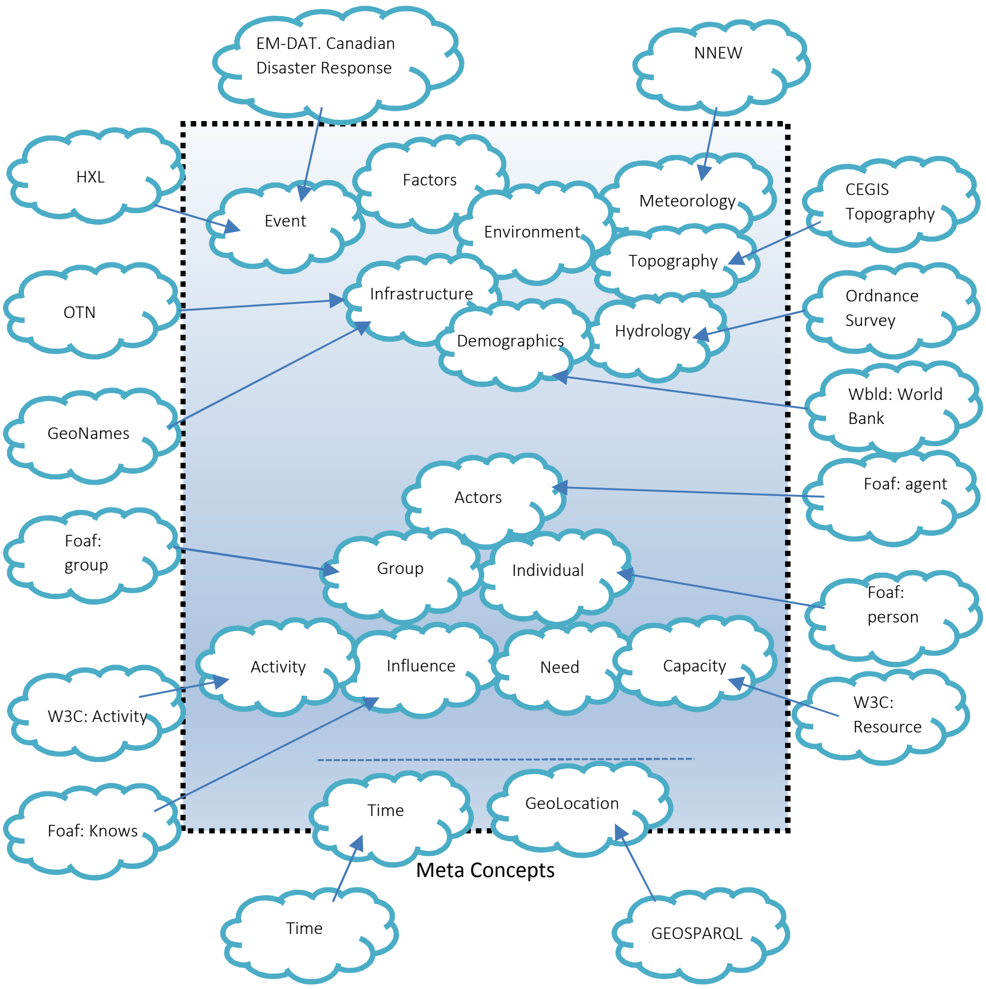


Figure 4: Reconciliation of Existing Ontologies with COBACORE Concepts

By reusing existing, 'well known' ontologies we automatically build extensibility into the model; as these ontologies mature so does the COBACORE model. Initial exploration reveals that many of the existing ontologies are incomplete, they are not domain specific so do not model the domain as accurately as the functionality requires. However, using semantic web technologies, we are able to extend and enhance these existing definitions to represent an ontology which is 'fit for purpose' in the COBACORE domain while, at the same time, keeping the interoperability advantages of reusing existing ontologies.

5.6 Semantic Architecture within the COBACORE Data Framework

The COBACORE data framework will be premised on a Service-Oriented Architecture (SOA) as advocated by Zhang et al. (2007), composing service provider, service broker, and service client (Figure 5). The OGC WFSs are used to publish feature-level data from heterogeneous Databases which are then mapped to OWL (Web Ontology Language)[3] existing (and if necessary expanded) disaster management ontologies to provide a semantic-based view of the services, which span from the abstract description of capabilities of the services to the actual feature data contents. The ontology server is used to create mappings of equivalent or related classes and properties in the existing ontologies. It keeps the taxonomy of disaster response terminologies and maintains the consistency of different existing ontologies.

To provide flexible query capability and ease of operation the framework, will allow emergency responders to type query questions in easily understandable natural language (free text). The parser layer will then be is used to convert query questions into SPARQL queries. SPARQL is used to access an OWL knowledge base in order to extract reasoning and meaning from the multi-source COBACORE Data framework.

[3] Because OWL is based on Description Logics (DL), it is intended to use a DL-based reasoner and inference rules to collect a knowledge base for automatic service queries – rather than this having to be done manually. The inference feature is still not fully developed as a concept and will form much of the thinking around the continued development and sophistication of the prototype over the remainder of the project. See Proposed Future Work Section.

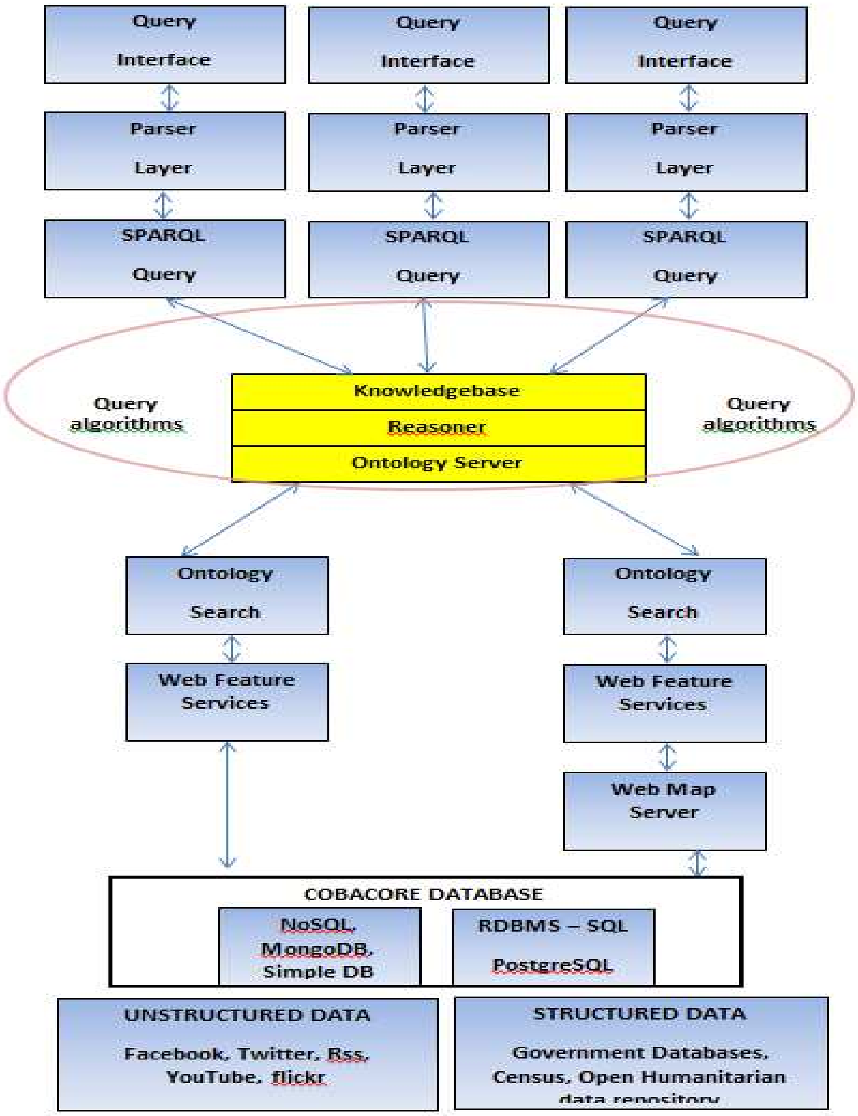
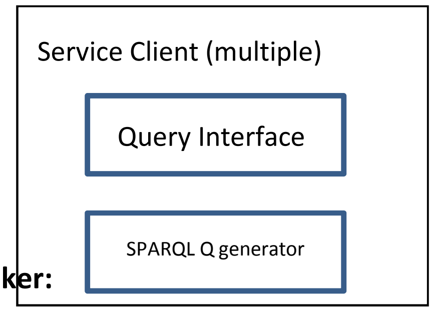


Figure 5: COBACORE Data Framework and Search Features/Functions

Service Clients:
(1) **Query Interface**: structured query in user's vocabulary
(2) **SPARQL Query Generator**: RDF query generated



Service Broker:

(1) **Ontology** in OWL: standard language

Knowledge base: a list of different data types

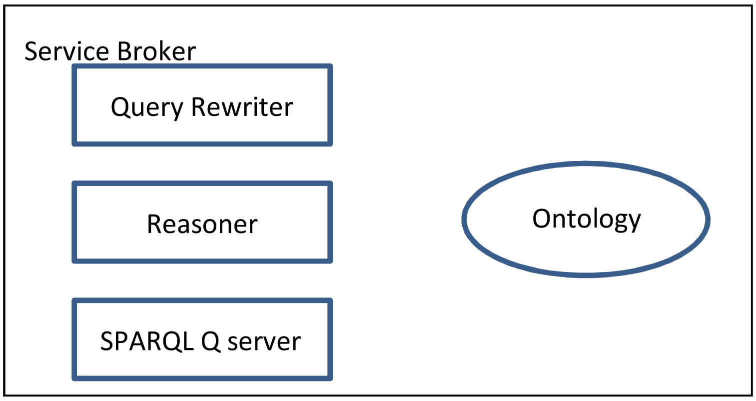
Shared vocabularies such as HXL

Thesaurus, which maps between shared vocabularies and domain/user specific Vocabularies

(2) **Query rewriter**: query rewritten to be domain specific

(3) **Reasoner**: mapping between vocabularies via reasoning. This module is called to rewrite query to make it domain specific.

(4) **SPARQL Query server**: (a) query broken down into domain specific sub queries using the knowledge base (b) query rewritten to be domain specific using the reasoner (c) send queries separately to different domain specific service providers

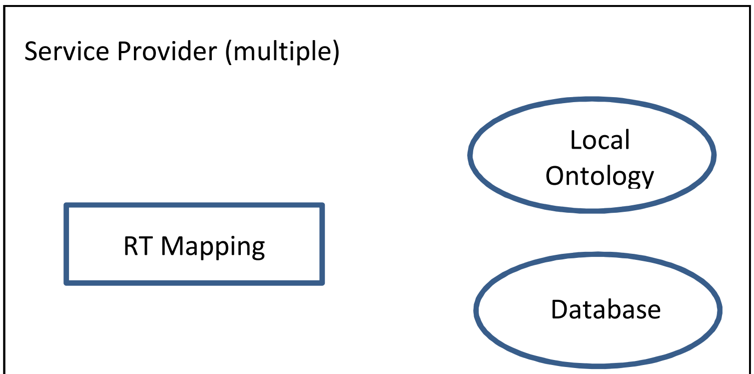


Service Providers:

(1) **Local (Domain) ontology** in OWL: vocabulary for a particular domain

(2) **RT Mapping**: it uses domain ontologies to transform a relational database into a set of RDF triples. We can use Ontop, a platform to query databases as Virtual RDF Graphs using SPARQL, to do the mapping between the relational databases records and the OWL ontologies. Ontop is a platform to query databases as Virtual RDF Graphs using SPARQL. It does the mapping between the relational databases records and the OWL ontologies.

(3) **Database**: some are static and some are dynamic; some are relational and some are triple-based.



The considered advantages of the COBACORE data framework are:

- the framework can recognize and represent the implicit and explicit meaning of the heterogeneous disaster response data content and can achieve data interoperability at the semantic level;
- the framework can understand the semantic meaning of a query, and can convert the natural language query into a WFS understandable query or free text reasoning on the part of end-users;
- the framework can give appropriate answers to natural language queries, even if key query terms are not included in data contents or metadata. With the proposed framework, disaster responders do not have to manually parse their ideas and process heterogeneous data systems. The ideology is akin to that of a system of systems where interoperability is paramount.

5.7 Comprehensiveness

The approach taken to validate the data model is to make sure that it encompasses both the identified data sources and features.

Identified Data Sources

Many of the types of data required for a typical COBCORE scenario were identified in D2.1 predicated on a snow crisis scenario within Northern Ireland in 2013. The final dataset composition will be continue to be validated and verified to ensure transferability (across jurisdictions and different forms of disaster) during the evaluation exercises being developed in WP5. Alongside the evaluation exercises consultation will be conducted with prospective end-users (professional responders, affected community) in order to determine the optimum data set composition in order to enhance situational awareness and improve data sharing. If the model were unable to adequately represent all of the sources identified by these ongoing exercises then the data model could not be described as 'comprehensive'. As a way to validate the comprehensiveness of the model, the following table shows how each identified data source can be represented by the data model.

	Data Source	**Applicable Concepts**
Rivers & Flood Plain	Inspire	Hydrology, GeoLocation
Accessibility	Inspire: Road Networks, Google Maps, Open Street Map	Infrastructure, GeoLocation

Utilities – Gas, Electric, Water	Inspire	Infrastructure, GeoLocation
Administrative Area	Inspire, Google Maps, OSM, GeoNames	Infrastructure, GeoLocation
Terrain / Elevation	Copernicus, Inspire (Ortho, Terrain), Google Maps	Environment, GeoLocation
Demographics	Inspire, NISRA	Environment, GeoLocation, Demographics
Socio-Economic	NISRA	Environment, Demographics, GeoLocation
Land Use Classification	Inspire, Open Street Map, Pointer	Environment, GeoLocation
Vulnerability - Age	Inspire, NISRA	Actors, Individual, GeoLocation, Group
Vulnerability - Disability	NISRA, Inspire	Actors, Individual, GeoLocation, Group
Reports	Media, RSS, GEORSS, Twitter, Facebook, Instagram	Event, GeoLocation
Shelter Locations	Inspire, PSNI	Capacity, GeoLocation
Injuries / Deaths	Media	Environment, Actors, Group, Individual, GeoLocation
Search & Rescue	PSNI, RAF, Coastguard	Capacity, Group, GeoLocation
Emergency Supplies	NIWS, Local Government Websites	Capacity, Group, GeoLocation
Gritter Locations	DOE	Capacity, Group, GeoLocation
Emergency Services	NISRA	Capacity,

Locations		GeoLocation, Group
Road Closures	PSNI, RSS, social media, media	Infrastructure, GeoLocation
Incident Lines	DOE, PSNI	Infrastructure, Capacity, Actors
Responders location	Social Media, Media	Actor, Groups, GeoLocation
Weather Forecasts	Met Office	Meteorology

Table 6 Data sources and applicable concepts

No Regrets

To validate the data model, it must also be shown to be comprehensive enough to encompass all the required functional features (As detailed in section 3.0). The table below describes how all the identified concepts related to the 'no regrets' features.

No Regrets Feature	Applicable Concepts
Needs registration: a feature for Affected Community members to state their needs (creation of a needs profile – type, details, dependencies, time, severity, etc).	Need, Person, Group, Agent, Time, GeoLocation
Needs overview: an overview of the stated needs, projected onto a map, or in charts (a dashboard)	Need, Person, Group, Agent, Time, GeoLocation
Baseline overview: GIS based view of baseline information of the affected area (also known as 'a situation overview' – thematically organised information about the built environment, social environment, infrastructures, environmental aspects, and so on)	Factors, Physical, Demographics, Time, GeoLocation
Plan/activity/information sharing: facility to share plans, requests and information (mainly for professional community parties to share towards affected community members – based on need, location, or other profile elements).	Activity, Capacity/Resource, Person, Group, Time, GeoLocation
Capacity registration: a feature for responding	Capacity, Person, Group, Time,

community groups (and perhaps professional groups) to register their capacities (creation of a capacity profile – type, area, timeframe, costs, etc).	GeoLocation
Needs & capacity overview: a feature to display/browse needs and capacities that have been registered in the platform.	Need, Capacity, Person, Group, Time, GeoLocation
Needs & capacity matching feature: a feature to link a specified need with a specified capacity (either from the perspective of an AC member or a RC member – might be something simple as sending out a request for contact, or some other form of signal)	Need, Capacity, Person, Group, Influence, Time, GeoLocation
Instruction/document sharing feature: a feature for a PC member to share instructions, knowledge or other training material to specific RC members (a.k.a. the COBA-school)	Activity, Capacity/Resource, Person, Group
Capacity/activity/organisation overview: a feature to shows registered capacities and groups (organisation/capacity awareness)	Activity, Capacity/Resource, Person, Group, Time, GeoLocation
Activity suggestion feature: a feature that suggest activities that registered/registering RC groups might be interested in ('nudging' – e.g. if I register myself with a desire/capacity to help out with activity type X in area Y, then the platform would suggest fitting activities, that fall under the same type in that area).	Need, Capacity, Person, Group, Influence, Time, GeoLocation

Table 7 No Regrets Relationships to Concepts

5.8 Framework Validation and Verification

Validation and verification is the process of checking that a system or framework meets specifications and that it fulfils its intended purpose. These are critical components of any quality data management system. Validation checks that the design satisfies or fits the intended usage and verification is the process of evaluating the framework to determine whether the architecture of a given development phase satisfy the conditions imposed at the start of that phase.

The most powerful tool for supporting asymmetric knowledge and in order to enhance the capacity of end-users in order to optimize the benefits afforded by the COBACORE

data framework it is envisaged that periodic exercises will be conducted throughout the duration of the COBACORE prototype development. Whilst high-fidelity evaluations including scenario based field operations offer greatest value and insight in terms of the refinement of the data framework even low level fidelity exercises and continuous engagement with prospective end-users will help to develop a deeper appreciation and understanding of the needs and expectations of end-users in order to ensure that COBACORE is conducive to prospective end-user expectations and adds sufficient value to the enhancement of information flow over all phase of the disaster cycle to ultimately warrant adoption within the confines of a real life disaster response operation.

6.0 CONCLUSION

The development features and hierarchal structures pertaining to the COBACORE data framework detailed in this document have been compiled premised on the underpinning data requirements and operational features necessitated to facilitate conformance with the No-Regrets Profile (NRP) developed in WP3. The COBACORE data framework will nonetheless change, and have the ability to change, over several iterations relative to required features and functions and in light of findings derived from both the intermediate evaluation exercise and the final evaluation exercise designed in WP5. Presently, the COBACORE data framework has been developed into a conceptual prototype for application. The prototype will continue to be refined over the next 12 months (D2.3) to ensure interoperability and interface compliance with the platform being developed in WP4 and to ensure operational robustness and compatibility relative to feedback from prospective end-users encompassing professional responders and affected communities.

Ultimately the 'success' of the data framework will be determined by:

- Its ability to embrace the 'no regrets' features detailed in chapter three of this document.
- Its ability to be 'state of the art' by being semantically interlinked with a wide range of existing data sources deemed relevant to crisis response management by prospective end-users.
- Its ability to be used as the building block for WP3, WP4 and WP5. In order for WP3, WP4 and WP5 to move forward they require the data framework to be in a format which is usable for the concept and platform development activities that are to be delivered. The main output from this section will be a conceptual data model which will adequately facilitate the creation of use-case diagrams in line with function and feature requirements (for WP3) and a physical data model which is fully interface compatible with the COBACORE End-User platform (WP4). Furthermore the framework must be developed to an operational standard conducive with the intermediate and final evaluation exercises (WP5) as well as being suitably robust and 'useable' to attain credibility and viability as a 'proof of concept'.

7.0 REFERENCES

Babitski,G., Bergweiler, S., Hoffmann, J. Sch¨on, D., Stasch, C. and Walkowski, A. (2010) Ontology-based Integration of SensorWeb Services in Disaster Management.

Bénaben, F. Chihab Hanachi, C. Lauras, M., Couget, P. and Chapurlat, V. (2008) *A Metamodel and its Ontology to Guide Crisis Characterization and its Collaborative Management. Proceedings of the 5th International ISCRAM Conference – Washington, DC, USA, May 2008.*

Bakillah, M.,Mostafavi M. A (2007) A Real Time Semantic Interoperability Framework For Ad Hoc Network Of Geospatial Databases: Disaster Management Context.

Bharosa, N., Lee, J. and Janssen, M. (2009) Challaenges and obstacles in sharing and co-ordinating information during multi-agency disaster response: Propositions from Field Exercises. Inf Syst Front (2010), 12:49-65.

Chen-Huei, C., Fatemeh. Z., Huimin, Z. (2011) *Ontology for Developing Web Sites for Natural Disaster Management: Methodology and Implementation.* IEEE Transactions on Systems, Man, and Cybernetics, Part A 41(1): 50-62.

Chuang, A. W.-C.-R. (2010). An Event Ontology An Event Ontologyfor Crisis-Disaster Information Crisis. *PNC.* Hong Kong.

Fowler, M. W. (2011). *Developing High Quality Data Models. The European Process Industries STEP Technical Liaison Executive (EPISTLE).*

Keßler, C. a. (2009). The Humanitarian eXchange Language: Coordinating disaster response with semantic web technologies. *Semantic Web*, 1-5.

Lan, H.X, Martin., C.D, Froese, C.R, Kim, T.H, Morgand Chowdhury S., 2009, A web-based GIS for managing and assessing landslide data for thetown of Peace River, Canada, Nat. Hazards Earth Syst. Sci., 9,pp 1433–1443.

Limbu, M. K. (2012). *Ntegration of Crowdsourced Information With Traditional Crisis And Disaster Management Information Using Linked Data. I*

Liu, S. D. (2013). Ontologies for Crisis Management: A Review of State of the Art in Ontology Design and Usability. *ISCRAM.*

Noy, N. F. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology.* Chicago.

Xu, W. a. (2007). Ontologies for disaster management response. *Geomatics Solutions for Disaster Management*, 185-200.

Zhang, C. T. (2010). Automatic search of geospatial features for disaster and emergency management. *International Journal of Applied Earth Observation and Geoinformation*, 409-418.

https://www.youtube.com/watch?v=fQUCCk0g-_c

http://www.gdmc.nl/zlatanova/thesis/html/refer/ps/WX_SZ_2007.pdf

http://www.slideshare.net/andreasinica/20101203-an-event-ontology-for-disaster-informationhuang-and-chuang.

List of Acronyms

A

ANSI - An abbreviation for American National Standards Institute. ANSI standards have been established for many elements of computer systems to aid research and development. The existence of standards allows designers to develop general solutions to common problems.

Application Programming Interface (API) - specifies how software components should interact with each other. In practice, most often an API is a document that includes specifications for routines, data structures, object classes, and variables.

B

Base Map - Spatial data sets that provide the background upon which more specific thematic data is overlaid and analysed.

Big Data–a large volume of both structured and unstructured data that is so large it is difficult to process using traditional database and software techniques.

C

Cache – is a place to store something temporarily, when looking at a web-page the requested files are automatically saved on your hard disk in a cache so when you return to the page the browser will get the files from your hard disk rather than the original server.

Cloud –Commonly refers to network-based services which appear to be provided by real server hardware, which in fact are served up by virtual hardware. Such virtual servers do not physically exist and can therefore be moved around and scaled up (or down) on the fly without affecting the end user - arguably, rather like a cloud.

Common Operational Dataset (COD) – are critical datasets that are used to support the work of humanitarian actors across multiple sectors. They are considered the standard for the humanitarian community and should represent the best available datasets for each theme.

Conceptual Data Model - A conceptual data model is a summary-level data model that is most often used on strategic data projects. It typically describes an entire enterprise. Due to its highly abstract nature, it may be referred to as a conceptual model.

Conceptual Schema - This is a high-level description of a business informational needs. It typically includes only the main concepts and the main relationship among them. Typically this is a first-cut model, with insufficient detail to build an actual database.

Coverage – A data model for storing geographic features, a coverage stores a set of thematically associated data considered to be a unit. Features are stored as both primary features (points, arcs, polygons) and secondary features (ticks, links, annotation).

Creative Commons - is a non-profit organization headquartered in Mountain View, California, United States devoted to expanding the range of creative works available for others to build upon legally and to share.

Crowdsourcing–is a type of participative, normally online, activity in which an individual, institution, company proposes to a group of individuals with a varying knowledge, heterogeneity and number voluntary undertake a task. Often used to subdivide tedious work

or to fund-raise start-up companies and charities, this process can occur both online and offline

D

Data Harmonisation – the iterative process of capturing, defining, analysing and reconciling government information requirements, and data standardization as the mapping of this simplified data to international standards.

Dashboard – provides an 'at a glance' view of key performance indicators relevant to a particular objective or business process.

Disparate Data–is data from any number of sources, largely unknown and unlimited and in many different formats. Disparate Data are heterogeneous data. They are neither similar nor can be easily integrated with an organisations database management system.

F

Features – The collective term for individual objects shown on map layers in their correct locations. A 'Feature' could be a building, an area, a street, a road or postcode etc

Framework (Software) - An information architecture, a reusable software template, or skeleton, from which key enabling and supporting services can be selected, configured and integrated with application code. A Framework is a reusable software platform to develop applications, products and solutions.

Free Open Source Software (FOSS)– is computer software which can be classified as both free software and open source software. Therefore anyone is licensed to use, copy, study and change the software in a way and the source code is openly shared to encourage people to improve the functionality and design of the software.

Fundamental Operational Datasets (FOD) – datasets that are relevant to a humanitarian operation, but are more specific to a particular sector or otherwise do not fit into one of the COD themes.

G

Gazetteer – geographic dictionary used in conjunction with a map or atlas, will typically contain information concerning the geographic makeup, social statistics and physical features of a country, region or continent. Content of the gazetteer can also include the location and dimensions of the subject.

Geo-coding – The process of adding location information to a file or database so its objects can be displayed on a map, for example calculating the x,y location of properties based on their address.

Geographic Information System (GIS) – The entirety of a system used for processing location information including data, hardware, software, people and managed services.

Geography Mark-up Language (GML) – OpenGIS XML-based language for describing and encoding geospatial information. An application of XML, a specification developed by members of the Open GIS Consortium.

Geo server – is an open source software server written in Java that allows users to share and edit geospatial data using OpenGIS standards Web Feature Service (WFS), Web Coverage Service (WCS) and Web Map Service (WMS). Geo server can also publish data from any major spatial data source using open standards.

Geospatial Data – Digital data that represent the geographical location and characteristics of natural or man-made features, phenomena and boundaries of the Earth. Geospatial data represent abstractions of real-world entities, such as roads, buildings, vehicles, lakes, forests and countries. Geospatial data refers to such data in any format, including raster, vector, point, text, video, database records, etc.

GeoWebCache – is a java web application used to cache map tiles coming from a variety of sources such as the OpenGIS standard Web Map Service (WMS) in order to optimise map image delivery. GeoWebCache exists as a standalone application and as a built in extension to Geo server.

H

HTTP - Hypertext Transfer Protocol, the World Wide Web protocol for moving hypertext (HTML) files across the Internet.

Humanitarian eXchange Language (HXL) – an XML schema which allows humanitarian organisations to publish their data in a machine readable format. Once the data has been published in this open format then anyone can have access to it.

Humanitarian Free Open Source Software(H-FOSS) – refers to Free and Open Source Software created to be used in the support of Humanitarian response.

Humanitarian Open Streetmap Team (HOT)– coordinates the creation, production and distribution of free mapping resources to support humanitarian relief efforts in many places around the world. Their main objective is to promote the principles of open source and open data sharing to humanitarian response and support the growth of the OpenStreetMap project.

Hydrology - is the study of the movement, distribution, and quality of water on Earth and other planets, including the hydrologic cycle, water resources and environmental watershed sustainability

I

Infrastructure for Spatial Information in the European Community (INSPIRE) – Directive 2007/2/EC of the European Parliament. This seeks to ensure that the spatial data infrastructures of the member states are compatible and useable in an EC and a cross boundary context. INSPIRE will enable the sharing of environmental spatial information among public sector organisations and better facilitate public access to spatial information across Europe.

Information Model - An information model in software engineering is a representation of concepts and the relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse.

Internet Map Server (IMS) – server software which allows users to share, view and edit geospatial data via Open Geospatial Consortium (OGC) standards such as WMS, WFS and WCS.

K

Key Mark-up Language (KML) – an XML based language schema for expressing geographic features and visualisation on maps. KML was developed for use with Google Earth, in 2008 KML became an international standard of the **Open Geospatial Consortium**.

L

LIDAR – Light Detection and Ranging is an optical remote sensing technology that can measure the distance to, or other properties of a target.

M

Map - A two-dimensional visual portrayal of geospatial data. A map is not the data itself.

Metadata – refers to data about data and is used in two fundamentally different concepts

- Structured Metadata – design and specification of data structures
- Descriptive Metadata –describes a resource for purposes **Micro tasking** – is the process of breaking a large project into tiny, well defined tasks. The resulting micro tasks can be performed independently, cannot be automated and can be done in a short period of time i.e. less than an hour.

Meteorology - is the interdisciplinary scientific study of the atmosphere.

MOAC - the Management of a Crisis Vocabulary, is a lightweight vocabulary aiming to provide terms to enable practitioners to relate different "things" in crisis management activities together as Linked Data.

N

NGO - Non-governmental organization

O

Ontology – In computer science and information science, an ontology formally represents knowledge as a hierarchy of concepts within a domain, using a shared vocabulary to denote the types, properties and interrelationships of those concepts.

Open Data – is the idea that data should be freely available to everyone to use and publish as they wish, without restrictions from copyright, patents or other methods of control. Open Data has gained popularity in recent years with the launch of open-data government initiatives such as Data.gov and Data.gov.uk

Open Geospatial Consortium (OpenGIS)– an international industry consortium of 475 companies, government agencies and universities participating in a consensus process to publicly available interface standards. The standards enable developers to make complex spatial information and services available and useful to all kinds of applications.

Open Source –is computer software with its source code made available where the developer provides users the right to study, change and distribute the software at no cost to anyone for any purpose.

Open Street Map – is an initiative to create and provide free geographic data such as street maps, to anyone.

P

PostGIS – provides spatial objects / Features for the PostgreSQL database, allowing storage and querying of geographic information.

PostgreSQL– Powerful Open-Source Object-Relational DBMS, with a strong emphasis on extensibility and implements the majority of the SQL:2008 standard. PostgreSQL provides spatial extensions, PostGIS, to allow database operations to be performed using geometry as well as tabular data.

Projection – is the transformation of the latitudes and longitudes on a sphere into the locations on a flat plane i.e. a map.

R

Raster – is the method for storing spatial data that involves assigning a value to each dot in a large matrix e.g. Aerial photographs and satellite images

Resource Description Framework (**RDF**) - has come to be used as a general method for conceptual description or modelling of information that is implemented in web resources, using a variety of syntax notations and data serialization format

Relational Data Base - data in such a way that it can be added to, and used independently of, all other data stored in the database. Users can query a relational database without knowing how the information has been organised.

RSS – is a family of web feed formats used to publish frequently updated works such as blogs, news headlines in a standardised format.

S

Short Message Service (**SMS**) - is a text messaging service component of phone, web, or mobile communication systems, using standardized communications protocols that allow the exchange of short text messages between fixed line or mobile phone devices.

Social Media – the means of interactions among people in which they can create, share and exchange information and ideas in virtual communities and networks i.e. Facebook, twitter, YouTube, Flickr, etc.

SPARQL - is an RDF query language, that is, a query language for databases, able to retrieve and manipulate data stored in Resource Description Framework format.

SQL:2008 – is the sixth version of the ISO and ANSI standard for SQL database query language.

Structured Data – refers to data that is organized in a structure, the most common form of structured data is a database where specific information is stored based on methodology of columns and rows.

Structured Query Language (SQL) - is a standard interactive and programming language for getting information from and updating a database.

Styled Layer Descriptor (SLD) – is a WMS user defined XML encoding standard to allow symbolization and colouring of geographic features and coverage data.

T

Thematic Map – a map showing, by colour or pattern, the distribution of a single phenomenon.

Topography - is a field of planetary science comprising the study of surface shape and features of the Earth and other observable astronomical objects including planets, moons, and asteroids. It is also the description of such surface shapes and features (especially their depiction in maps). The topography of an area could also mean the surface shape and features themselves.

U

Unstructured Data – is information that does not have a pre-defined data model or is not organised in a pre-defined manner. Unstructured data is normally predominantly text, but may contain data such as dates, numbers and facts. Conversely, structured data refers to *data that is identifiable* because it is organized in a structure. The most common form of structured data is a database where specific information is stored based on a methodology of columns and rows

W

Web Coverage Service (WCS) – define services on coverage and supports the retrieval of geospatial data as “coverage”, the WCS supplies information in forms that are useful for client-side rendering, similar to the WFS.

Web Feature Service (WFS) - refers to the sending and receiving of geospatial data via the internet. An important distinction between WFS and WMS is that the WMS displays the geographic information after it has been rendered as a digital image, whereas WFS sends the “source code” of the map.

Web Mapping - Dynamic query, access, processing, combination and portrayal of different types of spatial information over the Web.

Web Map Service (WMS) – OpenGIS standard protocol for serving geo-referenced map tiles over the internet. Provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. A WMS request defines the geographic layers and area of interest to be processed. The response to the request will be one or more geo-registered map images, returned as a JPEG or PNG that can be displayed in a web browser application. The interface also supports the ability to style the returned images using the Styled Layer Description standard.

Web Services - are self-contained, self-describing, modular applications that can be published, located, and invoked across the Web. Web services perform functions that can be anything from simple requests to complicated business processes. Once a Web service is deployed, other applications (and other Web services) can discover and invoke the deployed service.

Widget–a reusable element of a Graphical User Interface, GUI

X

Extensible Mark-up Language (**XML**) - is a mark-up language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.