



Technological University Dublin ARROW@TU Dublin

Articles

School of Electrical and Electronic Engineering

2018-2

Towards a Digital Earth: Using Archetypes to Enable Knowledge Interoperability within Geo-Observational Sensor Systems Design

Paul Stacey Institute of Technology, Blanchardstown, paul.stacey@tudublin.ie

Damon Berry Technological University Dublin, damon.berry@tudublin.ie

Follow this and additional works at: https://arrow.tudublin.ie/engscheleart2

Recommended Citation

Stacey, p. & Berry, D. (2017) Towards a Digital Earth: Using Archetypes to enable Knowledge Interoperability within Geo-Observational Sensor Systems Design. *Earth Science Informatics* (2018). https://doi.org/10.1007/s12145-018-0340-z

This Article is brought to you for free and open access by the School of Electrical and Electronic Engineering at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie.



This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License



This is a pre-print of an article published in the Journal of Earth Science Informatics. The final authenticated version is available online at: <u>https://doi.org/10.1007/s12145-018-0340-z</u>

Towards a Digital Earth: Using Archetypes to enable Knowledge Interoperability within Geo-Observational Sensor Systems Design

Paul Stacey^{a,b,*} Damon Berry^b

^{*a*}Institute of Technology Blanchardstown, Blanchardstown Road North, Dublin 15, IRELAND ^{*b*}Dublin Institute of Technology, Kevin Street, Dublin 8, IRELAND

*Corresponding author: email: <u>paul.stacey@itb.ie</u>, phone: +353 (0) 87 9937988, ORCID: 0000-0002-4532-2981

Keywords

O&M; Two-level modelling; Archetypes; Digital Earth; Interoperable Knowledge Representation

Abstract

Earth System Science (ESS) observational data are often inadequately semantically enriched by geoobservational information systems in order to capture the true meaning of the associated data sets. Data models underpinning these information systems are often too rigid in their data representation to allow for the ever-changing and evolving nature of ESS domain concepts. This impoverished approach to observational data representation reduces the ability of multi-disciplinary practitioners to share information in a computable way.

Object oriented techniques typically employed to model data in a complex domain (with evolving domain concepts) can unnecessarily exclude domain specialists from the design process, invariably leading to a mismatch between the needs of the domain specialists, and how the concepts are modelled. In many cases, an over simplification of the domain concept is captured by the computer scientist.

This paper proposes that *two-level modelling* methodologies developed by Health Informaticians to tackle similar problems of specific domain use-case knowledge modelling can be re-used within ESS Informatics. A proposed methodology to re-use two-level modelling within geo-observational sensor systems is described. We show how the Open Geospatial Consortium's (OGC) Observations & Measurements (O&M) standard can act as a pragmatic solution for a stable reference-model (necessary for two-level modelling), and upon which more volatile domain specific concepts can be defined and managed using *archetypes*. A use-case is presented, followed by a worked example showing the implementation methodology and considerations leading to an O&M based, two-level modelling design approach, to realise semantically rich and interoperable Earth System Science based geo-observational sensor systems.

Introduction

Earth System Science (ESS) portrays the Earth as an intricate network of interrelated entities (NASA 1986). In ESS, Earth is viewed as a complex, evolving planet that is characterised by continuously interacting physical and biological change (Mackenzie 2010). Changes within Earth's processes occur across a wide range of geospatial and temporal scales. Understanding the extent of change between interrelated Earth processes in terms of time, space and scale is important for making higher-level decisions (for example relating to human populations and biological related industries such as agri/aquaculture). Reflecting the multi-faceted nature of the system that they study, Earth Systems Domain Scientists comprise a variety of domain specialists. ESS Scientists have a need to combine data from many sources and in various ways (ideally computable ways) to synthesize new understanding and document new knowledge (Di et al. 2002).

In-situ remote sensor deployments and satellite based Earth Observation (EO) systems monitoring environmental phenomena are two important sources of computable data for Earth Scientists (Hart, Martinez 2006). Observation systems are often built in isolation, and the data representations and associated documentation systems where they exist, are not adequately designed for secondary use and higher order knowledge generation.

In 1998 US Vice President Al Gore set out his vision for what he termed "Digital Earth" (Gore 1998). Gore's vision was a challenge to a diverse global community to enable the increasing amount of raw geospatial data to be combined and processed into understandable information. To achieve the Digital Earth vision, Gore highlighted the need to break into the vast silos of data and make these data sets accessible and suitable for secondary use.

In recent times, many countries and jurisdictions have established their own EO systems (Westerbeeke et al. 2006) (GEO 2016). NASA's Earth Observing System Clearing House (ECHO) (Pfister et al., 2001; ECHO 2005) and the European Earth observation programme Copernicus (EO/Copernicus 2016) are good examples of how heterogeneous systems are being developed. In addition to these relatively monolithic satellite-based remote sensing systems, there is an even larger number of heterogeneous insitu remote sensing systems for capturing and publishing useful data. In general, there are a plethora of heterogeneous monitoring systems with different access protocols, syntax, data types, identifiers, coding systems and metadata models; that in their current state do not provide any clear mechanism for interoperability even at the most fundamental data representation level. Cross-community sharing of computable information is therefore difficult to achieve in practice. Barriers to interoperability of this type within ESS Informatics means that ESS domain specialists cannot fully exploit the data that may be available and this compounds the problem of vast data silos referred to in Gore's vision.

The challenges of interoperability are well documented and are core elements of many research agendas including Geographical Information Science (Yuan et al. 2005). Much of the work done to date within the Information Science community has been to enable interoperability through standardisation, particularly at the syntactic level. However, Goodchild argues that:

"Standards have the effect of codifying and constraining, whereas geographic information is evolving rapidly, demanding a much more flexible approach to metadata that reflects changing needs and expanding context." (Goodchild 2006)

Goodchild's statement is valid for all complex, evolving domains where domain concept models also need to reflect that evolution; and traditional meta-data modelling techniques are employed. Grossner et al. (2008) refer to this system evolution requirement as *extensibility* and list it as an essential component for a Digital Earth system. Other essential components listed in the context of a Digital Earth system are semantic and ontology-based data models and *object-level metadata*. Object-level metadata refers to the need to distinguish and manage observational data and derived knowledge.

Recently, Craglia et al. (2012) set out a renewed vision for a Digital Earth leading to 2020, and highlighted several grand challenges. Solutions to some of these challenges are emerging through semantic interoperability efforts for geo-observational data systems. Standards such as the Open Geospatial Consortium's (OGC) Observations & Measurements (O&M) standard (Cox 2006) and ontology alignment efforts (Fernandez, Susel, 2013); standard meta-data models and controlled vocabularies (GEMET 2012) with the linking of instance data to ontological concepts and terminologies for semantic interoperability (Wölger et al. 2011) (Leadbetter et al. 2016) using vocabularies such as SeaDataNet (Schaap, Lowery 2010.) available on the NERC vocabulary servers (Leadbetter, Lowry, Clements 2012) are all helping to realise the Digital Earth vision. However, this paper argues that current solutions do no not go far enough. Particular use-case knowledge and understanding is not being provided for in an interoperable and computable way. Domain concepts are constrained early on and are too inflexible for practitioner's needs. The practice of constraining knowledge at an early acquisition stage is inherent in object-oriented techniques (Boegl, Adlassnig et al. 2004) and leads to impoverished concept descriptions, unrecorded knowledge (Gahegan, Pike 2006) and creeping system obsolescence (Beale 2002). Knowledge sharing can be maximised across an interdisciplinary super-domain (such as ESS) by empowering domain specialists with the ability to model domain concepts in a computable way themselves, and by allowing for the evolution of the domain concepts within the model.

In recognition of the relatively slow pace of evolution of data standards, particularly those that normalise information models, and the problem that coded terms or ontologies alone are not sufficient to achieve semantic interoperability, for over 20 years, health informaticians have been developing a highly sophisticated approach to information modelling that has become known as *two-level modelling*. Two level models (Beale 2002) are designed to facilitate large-scale sharing of high quality, multifaceted, flexible and durable documentation.

In order to help realise the paradigmatic shift and enable the realisation of the dynamic Digital Earth framework called for by Craglia et al. (2012) the authors have investigated two-level modelling techniques as a possible solution to manage how information and knowledge concepts are modelled and managed (Stacey, Berry 2015). Diviacco and Leadbetter (2017) have also recognised the possible application of this technique within an Earth Systems Science context. The next section introduces the reader to these proposed solutions and highlights how the resulting information and documentation can be re-used within the Earth Systems Science domain to aid the development of ESS based geoobservational systems and progress the vision of a Digital Earth.

A Two-Level Modelling Approach

Traditional information system design tightly couples information and knowledge concepts. This coupling happens early on in the design process, at the point where object and data models are developed. Beale (2002) refers to these type of design methodologies as "single-level" models. Beale

argues that where the single-model approach is applied to information systems in a constantly changing environment, these systems become expensive and difficult to maintain. Beale also notes that these types of systems need to be replaced after a number of years. The reason for this is that domain concepts are hard-coded into the software. As the domain evolves and changes, the software becomes outdated and less useful.

Single-level systems have also been shown to have limited interoperability, as they may not adhere to a standardised formal model. Beale postulates the core issue for creeping obsolescence in single-level information systems is the constant evolution of the knowledge in a domain. Flexible design methodologies are needed to keep up with the non-static nature of the domain. Beale proposes a two-level solution.

In the two-level approach a traditional object model is still developed. This is referred to as the *"Reference Model"* (RM). Beale refers to this as the first-level model. The second-level model is where the formalism of the domain knowledge is captured. The separation of domain concepts can be organised as follows

- 1st level: This is the informational level and contains what are described as the non-volatile concepts required to be modelled for the system. It is a reduced set of classes that have a very abstract meaning, but nevertheless, have features to incorporate data types, terminology or ontology bindings. These concepts can be used as "building blocks" according to rules described in level 2.
- 2nd level: This level is the knowledge level where the concepts that will undergo evolution over time are captured and can be bound to ontologies as required. These concepts are volatile in nature and can evolve over time as knowledge evolves without "breaking" the system. This level is captured as a knowledge model using "*archetypes*" and an *Archetype Model (AM)*.

The separation of (recorded or documented) information and (generally applicable) knowledge in information systems design allows a more flexible representation of the domain knowledge (e.g. as part of a separate ontology). In a two level model, the reference model contains features to allow individual ontological terms to be "bound" dynamically to any point in the information model while keeping a rigorous formal definition of the data that is being recorded.

Archetypes

The capturing of non-volatile or stable concepts in the 1st level, or reference model, can be achieved using traditional conceptual modelling approaches. When a reference model has been developed, the challenge is then, how are the semantics of the reference model, or the knowledge concepts that have not been captured by the reference model to be defined and implemented?

Beale (2002) notes that knowledge level concepts are essentially constraints on the reference level concepts. As such, the knowledge level can be captured as a set of *constraint statements*. Here a set of constraint statements are referred to as an archetype. Constraints provide semantic relationships between elements based on knowledge. Using archetypes, an archetype-model can be developed that formalises the volatile knowledge concepts within the 2nd level of a two-level based information system (Figure 1).

A formal language Archetype Definition Language (ADL) (Beale 2007) for defining archetypes exists and is maintained by the openEHR foundation (Kalra, Beale, Heard 2005). ADL is used to constrain information models. ADL is not dependent on the reference information model, but is best suited to information models that are very generic in nature, and so in principle can be used for ESS modelling.

Two-level Modelling for Health Information Systems

There are a number of parallel attempts at two-level models in healthcare; openEHR and CEN/ISO 13606 (ISO/TC 215, 2006, 2008, 2009a,b) are currently the most notable. Both models employ ADL, and embed the following data quality enhancing features.

- A strong recognition that the model is intended for documentation of phenomena, rather than for producing a general model of reality (Beale 2003). The latter is the role of an ontology (Peirce, 1935) but in the healthcare community; this is not considered to be the same as documentation. As mentioned previously, ontological information is "bound" into the information model, which focuses on documentation.
- Use of commonly agreed identifiers and related mechanisms to uniquely identify phenomena that are the subject of documentation or provide context for the document.
- Use of an evolved temporal model and time-based data types to allow different aspects relating to time to be recorded accurately and satisfactorily.
- Embedded or "bound" ontological codes at appropriate points in the two level model for referring to commonly agreed concepts and terms.
- Employment of a general and reusable reference model, composed of building block concepts that can be used in many different documentation scenarios. These models are quite similar in intent to the OGC's O&M model. As previously noted, this reference model corresponds to the first of the "two levels". Figure 2 shows a simplified EN 13606 reference model.
- Development of a consensus-based library of archetypes.

Two level models and archetypes go beyond the idea of "recording measurements" to developing community-standardised "documentation" that is designed by consensus of the members of the community themselves. The process of developing these archetypes is a slow one, but the benefits are considered to be worth the great effort.

Two-level Modelling in non-Health Systems

The separation of information and knowledge in information systems design allows a more flexible representation of the knowledge on one level while keeping a rigorous formal definition of the data on the informational level. The two-level model allows for the development of more flexible and scalable systems. It encourages the future proofing of systems; "cradle to grave". The main goal of the two-level approach is that it acknowledges the reality of, and thus supports knowledge evolution within a given domain. These characteristics of an information system are not just desirable within Health but have wide applicability, especially within ESS informatics. Until recently, two-level modelling approaches have been the preserve of Health Informaticians. However, Lezcano et al. have shown how the semantic integration of sensor data with disaster management systems can be facilitated using a two-level modelling approach (Lezcano, Santos, Garcia-Barriocanl 2003). Stacey and Berry (2015) along with Diviacco and Leadbetter (2017) have noted the potential benefits of a two-level modelling approach for geo-observational systems. While proposing a migration of the two-level modelling approach from

Health to other domains, it is necessary to be cognisant of the differences that exist.

Within healthcare informatics, the primary subject of documentation is the patient. The prevailing consensus within health informatics is that the patient should remain the dominant subject of documentation for shared electronic health care documentation. This is a primary difference between healthcare informatics and geomatics. The subjects of documentation in geo-information and documentation in Earth System Science are diverse. Diviacco et al.'s (2015) work with *boundary objects* highlights this diversity, and shows the current efforts within the geo-sciences community to tackle automatic semantic and dynamic knowledge representation. Beaulieu et al (2016) also highlight the growing need and the current state-of-art in cyberinfrastructures to support collaborative processes and semantic communication amongst a diverse set of domain specialists. This automatic recording of information is "a matter of life and death" for the subject of documentation. This is usually not the case in geospatial measurement and documentation. The patient safety, quality of care issue imposes a strong need for rigour in healthcare, and a certain conservatism about changing processes and systems.

In translating two-level models from the healthcare domain to the ESS domain then, the following technical tasks would need to be performed:

- 1. Develop a generalised identity model.
- 2. Develop functioning binding to ESS coding.
- 3. Develop a reference model.
- 4. Development of two-level information communication and processing for resource constrained devices.
- 5. Formation of a suitable community of supporters.
- 6. Development of consensus-based ESS archetypes.

Of course, those technical tasks would need to be supported by the same vigorous type of communitywide engagement and dissemination that has characterised the adoption of two level models in the health domain.

The remaining sections of this paper detail the on-going work of translating two-level models to the ESS domain, specifically for use in geo-observational sensor systems design. Acknowledging that the OGC's O&M standard is fundamental to our discussion, we firstly examine O&M as a core component of any proposed solution.

Two-level Modelling use for Geo-Observation Sensor Systems

(Beale 2002) comments that one of the main problems with "standard" models is that "they embody no single point of view". Standard models do not deal with change very well and invariably implementations tend to wander to accommodate the peculiarities of the particular implementation.

O&M is part of the OGC's Sensor Web Enablement (SWE) architecture (Botts et al. 2008) and is a semi-structured model. However, as noted by Beale (2002), while semi-structured models are an improvement on standard models, they introduce additional problems; strong typing is typically lost, while the model is still partially concrete as assumptions about the information are made and encoded.

As interoperability at the knowledge level is a key requirement for future ESS observational Systems and the realisation of a Digital Earth system, it is proposed that stable domain concepts be captured using the OGC's O&M standard and concepts with the potential to evolve, or "volatile concepts", be captured using techniques derived from a two-level modelling approach (ultimately archetypes).

Archetypes provide a mechanism for avoiding the pitfalls of over-codification within a singular-model and offer more advantages over semi-structured models. It has been noted that a two-level model is explicitly designed to work alongside a common terminology (in the form of ontologies) to support the creation of archetypes. An ontology is an explicit specification, which formalises a *conceptualisation* of a body of knowledge, in some area of interest (Gruber 1983). Within ESS domains there are many advanced domain vocabularies and ontologies, examples include, the Semantic Sensor Network Ontology (SSNO) (Compton, Barnaghi et al. 2012) and Semantic Web for Earth and Environment Technology ontology (SWEET) (Raskin, Pan 2005), which may be used to provide semantic support to a two-level approach. Whereas, a reference model is a collection of coherent information models and should capture the stable non-volatile concepts within a domain, ontologies are typically organised into levels. Foundation concepts tend to not change over time and are used to produce more specialised domain concepts in *sub-ontologies*. Reference Models therefore should use concepts from upper-level ontologies or knowledge concepts from the foundation/principles level in a multi-level knowledge space (Figure 3).

Work done by Probst and Florian (2006) on developing an ontological representation of O&M allows us to assess O&M in terms of an ontological hierarchy. O&M as an ontology would not be classified an as an upper-level ontology in the strict sense (Cox 2015a) (Cox 2015b). Examining the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (Masolo 2003) UltraLite (which is an upper-level ontology) and the Semantic Sensor Network Ontology (SSNO) alignments highlights this. In fact, upon closer examination the definitions of *Observation* in both O&M and SSN show that definitions of Observation within these two ontologies are not semantically equivalent, but merely a close match (Figure 4). Despite O&M's ontological representation at a sub-ontology level, O&M is considered stable within its given domain. A pragmatic decision would be to consider O&M concepts as a principles ontological level from which content, organisational and storage concepts can be derived within a geo-observational sensor system (Figure 3).

Archetyped O&M

Ideally, any reference model should be formed using level-0 principles concepts (Figure 3). When proposing O&M as a valid reference model, one must question whether O&M represents level-0 principles. DOLCE UltraLite could serve as a level-0 principles concepts-based reference model for this purpose. In that scenario, O&M could act as a basis for discovering content level concepts. For this work, the stability of O&M concepts is assumed to be sufficient to act as a set of level-0 knowledge concepts and hence as a basis for a reference model. Another question arises here, what principle concepts represented as classes within an O&M aligned reference model can and should be archetyped? An observation consists of: phenomenon, location, value, time, producing sensor: or OM: Observed Property, OM:FeatureOfInterest.location, OM:Result. OM: Observation. phenomenon Time. OM: Procedure. These principle concepts can be used to construct content level domain concepts. ObservedProperty allows for content level domain concepts to be further defined e.g. Temperature. The authors have examined O&M's suitability as a reference model and found that O&M does not contain all the necessary base and container types or appropriate *aggregation patterns* that are required for two-level modelling. Keeping in mind to ensure the core O&M standard is maintained, while still enabling O&M's use as a reference model in a two-level modelling approach, a recommended augmentation of an O&M aligned reference model design pattern has been developed. The proposed augmented O&M model is presented next.

An Augmented O&M Model

Within the two level modelling approach, the only way in which data instances can be created is from the reference model. Any two-level modelling reference model must provide a means of representing entities that are not concretely modelled. Using a *compound/element pattern* (highlighted in green, Figure 2) within a reference model allows the creation of recursive aggregation of domain specific concept objects from the non-volatile concepts captured within the reference model.

A proposed augmented O&M model incorporating the necessary compound/element patterns to facilitate domain specific concept creation is shown in Figure 5. The compound/element pattern needed within any reference model, and noted in Figure 2 can be clearly seen in Figure 5 (also highlighted in green). *GeoData_Composition* represents a meaningful aggregation level. At this level within our representation, a basic flexible identity model is provided for (see IDENTITY_ABSTRACT in Figure 5). However, the question of a generalised identity model within ESS information systems remains an open question. Chen's (2016) work on generalised identity models for healthcare may provide a way forward within the ESS domain. The GeoData_Composition pattern provides a mechanism for domain practitioners to extend the model and create document level knowledge representation of specific use case domain concepts. Note that O&M currently represents a *model of reality*, the augmented O&M model provides for a *model of recording* in addition to the model of reality, this is a well-established design principle in health informatics (Beale 2003).

The structure of the document provides for three levels of meaningful aggregation from which concepts can be created at the necessary ontological levels i.e. *Storage*, *Organisational* and *Content* (Figure 6). The reference model itself captures the principle ontological level, or the stable concepts within the domain.

The Art of Archetyping

One of the core principles of the two-level modelling approach is that it should enable domain practitioners to capture specific domain knowledge concepts and to manage them as they evolve over time. The 1st level, or reference model, is still developed by Informaticians. The 2nd level, or the knowledge level, is developed by a mixed group of authors, that include the domain practitioners themselves and who now have greater influence on the evolution of models within a community environment.

Figure 7 shows the separation of the two levels and highlights the mapping of volatile concepts at the Storage, Organisational and Content ontological levels to the stable concepts within the domain at the principles ontological level. As noted previously, all data instances are of the reference model. However, these principle concepts are constrained at runtime using the knowledge model, or archetypes. Archetypes are constraint statements and an archetype model represents a rich knowledge level model of

domain concepts. *Information instances* are created at runtime from the reference model. These instances also adhere to the constraints defined within the archetype model (Figure 7).

Community development of Archetypes is a complex task that is performed by domain specialists. Within health a sophisticated framework of tools have evolved over the past number of years to facilitate the development, management and evolution of domain specific Archetypes (Sundvall et al. 2008) (Maldonado, Moner et al. 2009) (Chen, Klein 2007). Figure 7 shows a holistic but simplified view of the architecture necessary to support the use of archetypes within a geo-observational sensor based connected system.

Archetyping in an Earth Systems Science Domain

The authors have undertaken a proof-of-concept archetype development exercise using the proposed augmented O&M model (Figure 5). As this is a novel exercise, there are currently no existing tools to fully aid archetype development outside of the Health domain. However, there are a number of health informatics based tools that can be used to aid initial development in other domains such as ESS. The Biomedical Informatics Group at the ITACA Institute at the Universitat Politècnica de València have developed the LinkEHR platform (Maldonado, Moner et al. 2009). The LinkEHR platform includes an Archetyping Editor tool that allows for the development of Archetypes from any reference model. EHRFlex (Blobel et al. 2010) provides a flexible tool that may be used to further two-level modelling outside of the health domain. Currently EHRFlex only supports CEN/ISO 13606; support for any archetype-based standard is planned in the future. The OpenEHR Java reference implementation allows for further development of existing tools for non-health domains (Chen, Klein 2007). Using LinkEHR, the authors have demonstrated how a serialised XML form of the augmented O&M model can be used to develop archetypes. The LinkEHR editor provides a visual development tool for the creation of archetypes or reference model constraint statements. This visual approach to archetype development enables domain specialists, who may be non-technical or expert in information modelling, to produce the required content models. Once the serialised (XML for example) form of the reference model is available, the visual modelling tool approach can then be used by domain practitioners to meet their needs. Once domain modelling has been agreed upon using the visual interface, LinkEHR will output an ADL representation of the archetype model so that it may be machine readable within a supporting system (as shown in Figure 7). Listing 1 shows an example of an ADL representation of an archetype developed using the LinkEHR editor. In this particular use-case, the authors attempt to model a marine data buoy. The augmented O&M reference model is constrained for the particular use-case using the archetype model (i.e. the constraint statements).

It is important to note that archetypes or constraint statements for particular use-cases are agreed upon by the community, domain experts or practitioners using visual tools. It is this ability to derive community agreed standards through the consensus of empowered domain practitioners that offers the real benefits in terms of knowledge interoperability of systems. Archetype models evolve progressively and "naturally" as the community's knowledge and understanding of the domain advances. The community evolution of archetypes is in contrast to the more traditional development of information models and standards; which happens on a longer time cycle, in a more top-down approach. In any usecase, it is acknowledged that there is a need to have a general agreement on the basic structural elements of the information; this is the role of an O&M based reference model. However using two-level modelling it is also possible to acknowledge the need of specific practitioners within an ESS community to agree on specific data sets for particular purposes that can be easily changed with evolving requirements and understanding. In this simplified use-case that is nevertheless useful for the purposes of illustration, the ESS group or community could be an expert oceanic group or international research project wishing to share information & knowledge from a set of globally deployed data buoys.

The archetype shown in Listing 1 captures a storage level concept *Weekly_Buoy_Data*, which is a constraint on the principles level concept GeoData_Composition. As GeoData_Composition exists in the reference model, it is assumed that it is stable and generic to the point of being usable as a base concept for the derivation of all other information concepts.

```
archetype (adl version=2.0.5)
   TPOT-OM-GeoData Composition.Weekly_Buoy_Data.v1
concept
   [at0000]
language
   original language = <[ISO 639-1::en-ie]>
description
    original author = <["organisation"] = <"tPOT">>
definition
              omposition[at0000] matches {
                                                                        -- Weekly Buoy Data
    GeoI
        archetype node id existence matches {0..1} matches {*}
        details cardinality matches {1..*} matches{....}
        observation set cardinality matches {1..*} matches{
             DBSERVATION SET[at0002] matches {
                 details cardinality matches {1..*} matches {...}
                observation existence matches {1..1} cardinality matches {1..*} matches{
                     OBSERVATION[at0004]occurrences matches {1..*} matches {
                         observedproperty existence matches {1..1} matches {....}
                         featureofinterest existence matches {1..1} matches {....}
                         . . . .
                         results existence matches {1..1} matches {
                             RESULTS[at0007] cardinality matches {1..*} matches{
                                 result occurrences {1..1} matches{
                                     RESULT [at0008] matches{
                                          details cardinality matches {1..*} matches{
    }
ontology
    terminologies available = <....>
    term definitions = <
        ["en-ie"] = <
            items = <
                ["at0000"] = <
                    text = <"Marine Data Buoy Weekly Report">
                    description = <"Marine Data Buoy Weekly Report">
    constraint definitions = < ....>
    term binding = <
        ["nerc"] = <
            items =<
                ["at0005"] = <[nerc::TEMPPR01]>
                 . . . .
```

Listing 1. ADL snippet representation of an archetype developed using the LinkEHR editor

Listing 1 serves to illustrate how constraint statements are captured in a machine-readable format. ADL is used in this instance as the representation; however the archetype may be represented in any serializable format such as XML or JSON. Listing 1, shows that this archetype provides a metadata description for a record of Weekly_Buoy_Data, which is a specialisation of the reference-model based concept GeoData_Composition. TPOT-OM refers to the augmented O&M reference model developed by the *towards People Oriented Technologies* (tPOT) research group at the Dublin Institute of Technology, of which the GeoData_Composition is a member. Weekly_Buoy_Data is a domain specific volatile concept, defined here in the 2nd knowledge level. Weekly_Buoy_Data is assigned the reference [at0000] and is defined further in Listing 1. Also of note, is the possibility to bind concept definitions to ontologies.

Again emphasising that an archetype is developed using a community consensus approach, we can see in Listing 1 (which is a simplified version) the archetype model allows a community to agree and document the domain specific use-case implementation specialisms needed on top of the reference model. This ability to document specialisms in this way enables the efficient management of the evolution of any system using archetypes. Grossner et al. (2008) refer to this ability as the *extensibility requirement* of a Digital Earth system. To-date any variance or specialisation needed during use-case implementation of O&M, has not been managed in any structured way. Also missing from the specialisation of O&M is a well-established general community consensus mechanism (INSPIRE 2007) (Klein 2009). This resultant unmanaged variance in the implementation of standards, although often done for valid reasons, is a barrier to semantic interoperability of systems, especially at the knowledge level.

Archetypes and Information Objects "in the Wild"

For a system to capture, store and serve data that adheres to the community knowledge model, the archetype is applied as a set of constraints to guide the production of the information objects at run-time. The archetype describes the structure and detail of instantiated records of information. At run-time, archetypes may be represented in memory in an *archetype-enabled kernel*. Archetypes are intended to be maintained using a Web based management, review, validation and publishing library system. Communities of domain experts access and contribute to the archetype management system, taking part in the review and validation process.

For the purposes of illustration, let us assume that a community of domain experts have agreed on, and validated the archetype *TPOT-OM-Geo_Data_Element.Weekly_Buoy_Data.v1*. Any geo-observation system serving data to the community implements the reference model, i.e. all data objects are produced from the underlying RM. In this case, a system would use the augmented O&M reference model as the basis for data object instantiation. As the RM would be considered stable and not subject to change over time, the core system software does not need to change over time. At run-time, the system constraints the data objects based on the constraints defined in the archetypes. Constraining of RM based data objects takes place at run-time through the processing of the machine-readable ADL file, using the archetype-enabled kernel. As the needs and understanding of the community of domain experts changes and evolves over time, the two-level based system also adjusts how it creates the data objects based on the corresponding ADL file that encodes the archetype constraints.



Listing 2. JSON representation of an information instance. The resulting information instance from the compound/element patterns within the augmented O&M reference model (Figure 5) highlighted in green. GeoData_Composition is the realisation of the inclusion of a model of recording/documentation within the O&M based reference model. The archetype_node_id attribute is inherited from the LOCATABLE class in Figure 5 and allows bindings to occur between instance data and the AM.

Listing 2 above depicts how an information instance may be represented. The O&M JSON encoding OM-JSON (Cox 2015) is used as the basis for our example. OM-JSON provides a number of schemas where validation of specialisations of O&M may be performed. In the methodology proposed in this paper, the base O&M schema would still be captured (reference model) but the specialisations (volatile quasi-static concepts) would be captured in the archetype model. Validation is one of the primary runtime uses of archetypes. Archetype-validation tools and frameworks are currently available (Chen et al. 2008). Using this approach the full power of the community consensus approach and associated tools would be available to evolve and manage specialisations. The authors believe that this approach is in keeping with, and helps to realise the vision of the dynamic Digital Earth framework set out by Craglia et al. (2012). Upon examination of Listing 2, it is of note that there is some overhead associated with this approach. It is necessary to record which archetype was used for data construction, this appears as *archetype_node_id*. Archetypes themselves are separate from their data and need to be stored in an

accessible repository (Figure 8). We can see that there are actually three levels of information in our example, *wholly-static concepts* that are captured in the reference model, *quasi-static concepts* which are agreed in each of the archetypes and *dynamic data* or instance information. A pragmatic approach to managing the growing volume of the dynamic data instance shown in Listing 2, is to identify additional static information from the information instance of dynamic data that may reside in the archetype, and remove this from the information instance. In reality any observational system should only need to report timestamped DATA_VALUES and the particular archetype to which the instance is bound, and to which it may be validated against. Presentation layer applications using data instances can then reconstitute the semantically rich information using a similar knowledge framework as shown in Figure 8.

The challenges of constrained in-situ remote sensor systems are not considered at this point; it is assumed that the sensor system has the resources necessary to support the archetype-template runtime environment and associated kernel. Figure 8 shows a number of separate supporting systems. Development of a library of community-derived archetypes is supported by the online management system and archetypes are available through an online repository. For any specific use-case, the system builder and associated domain specialists use the necessary subset of archetypes available with the library. These archetypes are further specialised for the particular use-case and are combined to produce a set of Operational Templates (OPT) (Leslie 2008). The hypothetical in-situ remote sensor system shown in Figure 8 uses these OPTs locally to instantiate information instances, such as the one shown in in Listing 2. Information instances may then be transmitted to a supporting data-store for persistence.

As information-instances are created using the reference model which in this case adheres to the O&M specification, our observation system can now conceivably publish semantically rich, interoperable data and information, which evolves as the knowledge community evolves. As the system also adheres to a core standardised information model, such as O&M (in this case), syntactic interoperability is maintained, and it becomes a relatively rudimentary task to make observations available to an OGC standardised Sensor Observation Service (SOS) (Bröring et al. 2012).

As archetypes have a predictable structure, derived from the underlying reference model, enhanced querying can be achieved using the Archetype Query Language (AQL) (AQL 2015). AQL is a fusion of SQL, and XPath style paths, derived from the archetype. Archetype paths transcend the archetype in to the instance data, in the form of archetype node identifiers (Listing 2). This ensures conformance of archetype path structures as data nodes are constructed at runtime, and allows data nodes to be extracted using complex queries.

Discussion

The authors believe that the approach presented here can enable a wide and diverse set of domain experts (within the Earth System Science community) to contribute directly to the creation and evolution of content models, while ensuring the provision of high quality and accurate shareable knowledge amongst a diverse super-domain. These characteristics indicate that the solution is worth considering for realisation of a Digital Earth system. For example, the provision of Web based "social network" style management, review and publishing of Digital Earth Knowledge Artefacts in the form of archetypes, would foster greater semantic interoperability and fulfil the need for an *integrated authority list*, a feature that was listed as an essential component for a Digital Earth system (Grossner et al. 2008). This continuous process of model evolution is in contrast to the relatively lengthy renewal cycle of ISO

standards. Templates and OPTs offer additional flexibility outside of the community-agreed archetype model for local uses. This provides for situations where disparate domain expert groups may disagree, and can lead to archetype alignment issues as the approach matures within the domain. Again, as the development community is richly supported, techniques to overcome this potential for divergence are emerging (Bisbal and Berry 2009).

Software tools to realise a two-level modelling methodology are complex to implement. Currently no toolset exists that is readily useable outside of the health domain. There are number of barriers to clinical based tool reuse; for instance, clinical domain modelling tools assume a static singular identity model, that of the Patient. However, as the authors have demonstrated in this work, open source two-level modelling tools and components can be re-used to aid ESS-facing tool development.

A key differentiator of two-level modelling compared to other approaches is that is allows domain experts to be the primary drivers of Digital Earth Artefacts, while also ensuring that technical validity is maintained in one highly accessible and integrated process; leading to a key *extensibility* component in a Digital Earth system. This view has also been expressed by domain experts (Clinicians) in the health domain following work with archetypes (Garde et al. 2007). Also of note from health domain experiences of two level models are reports of reduced complexity of software and a greater focus on the realisation of useful applications (Chen et al. 2009); arising from a reduced demand for software model authoring tools. Any increased focus on the more convenient realisation of useful applications in ESS only further helps to realise the *functionality actions* provided by a Digital Earth system as compiled by Grossner et al. (2008).

Development of a mature, consensus-based repository of community derived archetypes is a non-trivial task, and requires established processes within any domain to ensure proper governance (Wollersheim et al. 2009) (Garde et al. 2007). However, with over 15 years of development experience, the technique is well supported by a strong theoretical and methodological framework. The true benefits of two-level modelling and archetypes are certainly realised when a large community consensus approach is employed, but Hoy et al. (2007) show how smaller local communities can also begin seeing early dividends from two-level modelling, without a large archetype repository. This offers a way forward for new domains in terms of a parallel introduction of the technique. Should two-level modelling take hold as a preferred mechanism for the development of standard content models, migration of valuable ESS legacy systems to an archetype based representation is possible due to the rich expressive power of archetypes (Chen et al. 2009). Adoption of a common format for content models leads to rapid and convenient installation and configuration of new metadata. Employment of two-level modelling techniques could potentially facilitate a nationally or conceivably an internationally standardised representation of all ESS content, as is alluded to in the health domain (Bernstein 2009). This approach is ultimately about facilitating the pooling of high quality data between Earth System Scientists and helping to develop critical Digital Earth based decision support systems.

Conclusion

Current interoperability efforts solve many problems with ESS data interoperability. However, interoperability at a knowledge level to date has not been fully realised. Two-level modelling provides the basis for achieving interoperable knowledge based systems through the provision of tried and tested methodologies and constructs developed for the Health Informatics domain. With additional design

patterns, O&M can serve as the basis for a two-level modelling reference model. Care must be taken in relation to the context of use of O&M. In certain circumstances, O&M concepts may be better employed in helping to realise content level archetypes, with upper ontologies such as DOLCE UltraLite informing the reference model. This needs further exploration. The evolution of two level models in the health informatics domain has been consensus-based and slow, with support from the community gradually gathering pace. Community consensus is an important element of the approach. The health informatics community has learned valuable lessons in this time. The two-level modelling experience should be of interest to Earth System Scientists who wish to share interoperable information that is trusted across measurement platforms and sub-domains.

A significant body of knowledge and experience has developed over the last two decades in health informatics, which provides many approaches that are useful for solving knowledge interoperability problems within Earth Systems Science. Many of the solutions developed using the two-level modelling approach have parallel applicability for helping to realise a Digital Earth system and accrue the associated benefits. Further exploration of the applicability of a two-level modelling approach to ESS Information systems through the development of appropriate software tools is recommended.

Acknowledgements

The authors would like thank Dr. Adam Leadbetter from the Marine Institute, Ireland, for review comments and constructive feedback.

List of abbreviations

ADL Archetype Description Language AM Archetype Model AQL Archetype Query Language CEN European Committee for Standardization (Comité Européen de Normalisation) DOLCE Descriptive Ontology for Linguistic and Cognitive Engineering ECHO Earth Observing System Clearing House EHR Electronic Healthcare Record **EO Earth Observation** ESS Earth System Science **GEMET** General Multilingual Environmental Thesaurus **INSPIRE** Infrastructure for Spatial Information in Europe ISO International Organization for Standardization JSON JavaScript Object Notation NASA National Aeronautics and Space Administration NERC Natural Environment Research Council **O&M** Observations and Measurements OGC Open Geospatial Consortium **OPT** Operational Templates **RM** Reference Model SOS Sensor Observation Service SSNO Semantic Sensor Network Ontology SWEET Semantic Web for Earth and Environment Technology **TC** Technical Committee tPOT Towards People Oriented Technologies XML Extensible Markup Language

References

- AQL, Archetype Query Language 1.0.3 (2015) Available: http://www.openehr.org/releases/QUERY/latest/docs/AQL/AQL.html
- EO/Copernicus overview (2016) Available: http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview3.
- ECHO, Earth Observing System Clearing House,(2005) Overview. Available: <u>http://www.echo.eos.nasa.gov/overview/</u>.
- Beale, T. and Heard, S. (2007) Archetype Definition Language, The openEHR Foundation, London.
- Beale, T. (2002) Archetypes: Constraint-based domain models for future-proof information systems, OOPSLA workshop on behavioural semantics.
- Beale, T. (2003) Towards Design Principles for Health Information Systems, International Journal of Medical Informatics.
- Beaulieu, S.E., Fox, P.A., Di Stefano, M. et al. (2016) Toward cyberinfrastructure to facilitate collaboration and reproducibility for marine integrated ecosystem assessments Earth Sci Inform. doi:10.1007/s12145-016-0280-4
- Bernstein, K., Tvede, I., Petersen, J. and Bredegaard, K., (2009) Can openEHR archetypes be used in a national context? The Danish archetype proof-of-concept project. In *MIE* (pp. 147-151).
- Bisbal, J. and Berry, D., 2009. Archetype Alignment-A Two-level Driven Semantic Matching Approach to Interoperability in the Clinical Domain. In HEALTHINF (pp. 216-221).
- Blobel, B., Moner, D., Hildebrand, C., Robles, M., (2010) Standardized and flexible health data management with an archetype driven EHR system (EHRflex). In Seamless Care, Safe Care: The Challenges of Interoperability and Patient Safety in Health Care: Proceedings of the EFMI Special Topic Conference, June 2-4, Reykjavik, Iceland (Vol. 155, p. 212). IOS Press.
- Boegl, K., Adlassnig, K., Hayashi, Y., Rothenfluh, T.E. & Leitich, H. (2004) Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system, Artificial Intelligence in Medicine, vol. 30, no. 1, pp. 1-26.
- Botts, M., Percivall, G., Reed, C. & Davidson, J. (2008) OGC® sensor web enablement: Overview and high level architecture in GeoSensor networks Springer, pp. 175-190.
- Bröring, A., Stasch, C. and Echterhoff, J., (2012) OGC sensor observation service interface standard. Open Geospatial Consortium Interface Standard, pp.12-006.
- Chen, R., Garde, S., Beale, T., Nyström, M., Karlsson, D., Klein, G.O. and Åhlfeldt, H., (2008) An

archetype-based testing framework. Studies in health technology and informatics, 136, p.401.

- Chen, R., Klein, G.O., Sundvall, E., Karlsson, D. and Åhlfeldt, H., (2009) Archetype-based conversion of EHR content models: pilot experience with a regional EHR system. BMC medical informatics and decision making, 9(1), p.33.
- Chen, R. & Klein, G. (2007) The openEHR Java reference implementation project, Studies in health technology and informatics, vol. 129, no. 1, pp. 58.
- Chen, X., (2016) A two-level identity model to support interoperability of identity information in electronic health record systems. Thesis submitted for the degree of Doctor of Philosophy to Dublin Institute of Technology.
- Council, N.A. & Earth System Sciences Committee (1986) Earth System Science Overview: A Program for Global Change, National Aeronautics and Space Administration.
- Cox, S. (2006) Observations and measurements, Open Geospatial Consortium Best Practices Document. Open Geospatial Consortium.
- Cox, S. (2015a) Ontology alignment is PROV-O good enough? Open Geospatial Consortium GeoMatics Summit, Boulder, CO.
- Cox, S. (2015b) Pitfalls in alignment of observation models resolved using PROV as an upper ontology. In AGU Fall Meeting Abstracts.
- Cox, S. and Taylor, P. (2015) OGC Observations and Measurements JSON implementation, Open Geospatial Consortium, Tech. Rep. Discussion Paper 15-100r1. [Online]. Available: http://www. opengis.net/doc/dp/om-json/
- Craglia M, de Bie K, Jackson D et al (2012) Digital earth 2020: towards the vision for the next decade. Int J Digit Earth 5:4–21
- Directive, I. (2007) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Published in the official Journal on the 25th April.
- Diviacco, P., De Cauwer, K., Leadbetter, A., Sorribas, J., Stojanov, Y., Busato, A. and Cova, A., (2015) Bridging semantically different paradigms in the field of marine acquisition event logging. Earth Science Informatics, 8(1), pp.135-146.
- Diviacco, P., Leadbetter, A., (2017) Balancing Formalization and Representation in Cross-Domain Data Management for Sustainable Development. Oceanographic and Marine Cross-Domain Data Management for Sustainable Development. IGI Global, 23-46.
- Fernandez, S., Marsa-Maestre, I., Velasco, J.R. & Alarcos, B. (2013) Ontology Alignment Architecture for Semantic Sensor Web Integration, Sensors, vol. 13, no. 9, pp. 12581-12604.

- Gahegan, M. & Pike, W. (2006) A situated knowledge representation of geographical information, Transactions in GIS, vol. 10, no. 5, pp. 727-749.
- Garde, S., Hovenga, E. J., Granz, J., Foozonkhah, S. and Heard, S. (2007). Managing archetypes for sustainable and semantically interoperable electronic health records. eJHI electronic Journal of Health Informatics 2(2): 10.
- Garde, S., Chen, R., Leslie, H., Beale, T., McNICOLL, I. and Heard, S., 2009. Archetype-based knowledge management for semantic interoperability of electronic health records. In MIE (pp. 1007-1011).
- GEMET, General Multilingual Environmental Thesaurus, (2012) Available: <u>http://www.eionet.europa.eu/gemet/</u>
- GEO, Group on Earth Observations, (2016) Available: https://www.earthobservations.org/index.php
- Goodchild, M.F. (1991) Just the facts, Political Geography Quarterly, vol. 10, no. 4, pp. 335-337.
- Goodwin, C. & Russomanno, D.J. (2006) An ontology-based sensor network prototype environment, Proceedings of the Fifth International Conference on Information Processing in Sensor Networks, pp. 1.
- Gore, A. (1998) The digital Earth: understanding our planet in the 21st century, Aust Surv 43:89–91
- Grossner, K.E., Goodchild, M.F. and Clarke, K.C., (2008) Defining a digital earth system. Transactions in GIS, 12(1), pp.145-160.
- Gruber, T.R., 1993. A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), pp.199-220.
- Hoy, D., Hardiker, N. R., McNicoll, I. T., Westwell, P. and Bryans, A. (2009). Collaborative development of clinical templates as a national resource. International Journal of Medical Informatics 78 Suppl 1: S95-100.
- ISO/TC215 (2006) Health Informatics-Electronic health record communication-Part1 : Reference model.
- ISO/TC215 (2008) Health Informatics-Electronic health record communication-Part2: Archetype interchange specification.
- ISO/TC215 (2009a) Health informatics-Electronic health record communication-Part 3: Reference archetypes and term lists.
- ISO/TC215 (2009b) Health Informatics-Electronic Health Record Communication-Part 4: Security.

- Kalra, D., Beale, T. & Heard, S. (2005) The openEHR foundation, Studies in health technology and informatics, vol. 115, pp. 153-173.
- Klien, E., Annoni, A. & Marchetti, P.G. (2009) The GIGAS project-an action in support to GEOSS, INSPIRE, and GMES.
- Leadbetter, A., Lowry, R. and Clements, D.O., (2012) April. The NERC Vocabulary Server: Version 2.0. In Geophysical Research Abstracts (Vol. 14).
- Leadbetter, A., Smyth, D., Fuller, R., O'Grady, E. and Shepherd, A. (2016) Where Big Data meets Linked Data: Applying standard data models to environmental data streams, 2016 IEEE International Conference on Big Data.
- Leslie, H., (2008) International developments in openEHR archetypes and templates. Health Information Management Journal, 37(1), pp.38-39
- Lezcano, L., Santos, L. & García-Barriocanal, E. (2013) Semantic Integration of Sensor Data and Disaster Management Systems: The Emergency Archetype Approach, International Journal of Distributed Sensor Networks, vol. 2013.
- Mackenzie, F.T. & Mackenzie, J.A. (2010) Our changing planet: an introduction to earth system science and global environmental change, 4th edn, Prentice Hall New Jersey.
- Maldonado, J.A., Moner, D., Boscá, D., Fernández-Breis, J.T., Angulo, C. & Robles, M. (2009) LinkEHR-Ed: A multi-reference model archetype editor based on formal semantics, International journal of medical informatics, vol. 78, no. 8, pp. 559-570.
- Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. and Schneider, L., (2003) Dolce: a descriptive ontology for linguistic and cognitive engineering. WonderWeb Project, Deliverable D17 v2, 1, pp.75-105.
- Michelsen, L., Pedersen, S.S., Tilma, H.B. and Andersen, S.K., (2005). Comparing different approaches to two-level modelling of electronic health records. Studies in health technology and informatics, 116, pp.113-118
- Peirce, C.S. (1935) Collected papers. Vol. VI. Scientific metaphysics (Ed. by C. Hartshorne and P. Weiss.).
- Pfister, R., Ullman, R., Wichmann, K. & Perkins, D.C. (2001) ECHO Responds to NASA's Earth Science User Community.
- Probst, F., Gordon, A. & Dornelas, I. (2006) OGC Discussion Paper: Ontology-based Representation of the OGC Observations and Measurements Model, Institute for Geoinformatics (ifgi).

Raskin, R.G. & Pan, M.J. (2005) Knowledge representation in the semantic web for Earth and

environmental terminology (SWEET), Computers & Geosciences, vol. 31, no. 9, pp. 1119-1125.

- Schaap, D.M. and Lowry, R.K., (2010. SeaDataNet–Pan-European infrastructure for marine and ocean data management: unified access to distributed data sets. International Journal of Digital Earth, 3(S1), pp.50-69
- Stacey, P., Berry, D. (2015) Applying Two-Level Modelling to Remote Sensor Systems Design to Enable Future Knowledge Generation. In IEEE YP Conference in Remote Sensing, Barcelona, Abstracts
- Sundvall, E., Qamar, R., Nyström, M., Forss, M., Petersson, H., Karlsson, D., Åhlfeldt, H. & Rector, A. (2008) Integration of tools for binding archetypes to SNOMED CT, BMC Medical Informatics and Decision Making, vol. 8, no. 1, pp. 1.
- Westerbeeke, H., Secretariaat, G. & Geneve, Z. (2006) Group on Earth Observations (GEO), RUIMTEVAART, vol. 55, no. 4, pp. 10.
- Wollersheim D., Sari A., Rahayu W. (2009) Archetype-Based Electronic Health Records: A Literature Review and Evaluation of Their Applicability to Health Data Interoperability and Access. Health Information Management Journal, Vol 38, Issue 2, pp. 7 - 17.
- Wölger, S., Siorpaes, K., Bürger, T., Simperl, E., Thaler, S. and Hofer, C., (2011) A survey on data interlinking methods. Technical report, Semantic Technology Institute, Innsbruck.
- Yuan, M. Mark, D M. Egenhofer, M J. and Peuquet, D J. (2005) A Research Agenda for Geographic Information Science, Extensions to geographic representation. In McMaster R B and Usery E I (eds) Boca Raton, FL, CRC Press: 129–56



Fig.1 Two-Level Model separation of stable concepts from volatile domain concepts



Fig. 2 A portion of the EN 13606 Reference Model. Compound/element patterns are highlighted in green



Fig. 3 Ontological levels. Within a two-level model, O&M as a reference model should only capture stable concepts i.e. at the principles level. They should be true for all instances and all use contexts. Typical of Upper Level Ontologies (Beale, 2002). Here we map higher-level domain concepts derived from O&M principle level concepts onto the Beale's (2002) multi-level knowledge space



Fig. 4 DUL, SSN, O&M Alignment



Fig. 5 Augmented O&M model. Compound/element patterns are highlighted in green.

Document: Geo-Observations Record



Fig. 6 Document Structure



Fig. 7 An Archetype Model (AM) is used to constrain Reference Object Model (ROM) instances at runtime



Fig. 8 Architecture of a proposed geo-observational in-situ sensor system to support the implementation of a two-level modelling approach.