An Ontology Derived from Heterogeneous Sustainability Indicator Set Documents

Lida Ghahremanloo[1], James A. Thom[1], Liam Magee[2]
[1]School of Computer Science and Information Technology
[2]School of Global Studies, Social Science and Planning
RMIT University
Melbourne, Australia
(lida.ghahremanloo,james.thom,liam.magee)@rmit.edu.au

ABSTRACT

We present an *ontology* to represent the key concepts of sustainability indicators that are increasingly being used to measure the economic, environmental and social properties of complex systems. There have been few efforts to represent multiple indicators formally, in spite of the fact that comparison of indicators and measurements across reporting contexts is a critical task. In this paper, we apply the METHONTOLOGY approach to guide the construction of two design candidates we term *Generic* and *Specific*. Of the two, the generic design is more abstract, with fewer classes and properties. Documents describing two indicator systems – the Global Reporting Initiative and the Organisation for Economic Co-operation and Development – are used in the design of both candidate ontologies. We then evaluate both ontology designs using the ROMEO approach, to calculate their level of coverage against the seen indicators, as well as against an unseen third indicator set (the United Nations Statistics Division). We also show that use of existing structured approaches like METHONTOLOGY and ROMEO can reduce ambiguity in ontology design and evaluation for domain-level ontologies. It is concluded that where an ontology needs to be designed for both seen and unseen indicator systems, a generic and reusable design is preferable.

Categories and Subject Descriptors

H.2 [**Database Management**]: Logical Design—*Data models*; I.2 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods—*Representations (procedural and rule-based)*; D.2 [**Software Engineering**]: Interoperability

Keywords

Ontology Engineering, Ontology Evaluation, Sustainability Indicators

1. INTRODUCTION

Since the publication of the Club of Rome's "The Limits to Growth" in the early 1970s, the sustainability of natural and social systems has become a pressing concern. This problem has become increasingly urgent, given the depletion of our natural stocks due to past and present economic development. To respond to such a challenge, it is vitally important to develop reliable and robust *indicators*, to measure the sustainability of our complex economic, environmental and social systems. These provide the tools to manage current and future development responsibly. In response, a number of indicator systems have been developed and are in use today. These include generalized reporting standards, such as those developed by the Global Reporting Initiative (GRI)[1], the Organization for Economic Co-operation and Development (OECD)[2] and the United Nations Statistics Division (UN Social Indicators)[3].

The number of indicator systems itself poses problems for would-be reporting organisations – how to find which indicators best represent the challenges of their specific context? To date, there have been few efforts to represent *multiple* indicator systems in a systematic way. In particular, we see considerable applications for the representation of indicators in a formal *ontology*. In computing disciplines, an ontology refers to a formal explicit specification of a shared conceptualisation for a domain of interests consisting of a set of *concepts*, *relationships* and *individuals* that are reused within enterprises [13]. Through use of such a formal construct, it is possible to develop a consistent definition of what an indicator is, and how it can be applied. This in turn would allow organisations to browse and review different kinds of indicators for different measurement applications, and to enable some degree of comparison and bench-marking between them. A further challenge exists in the design and evaluation of domain-level ontologies. Although in ontology research several approaches have been proposed for structuring the knowledge into different levels of abstraction [7, 21, 24], there is still considerable ambiguity involved in the processes of construction and evaluation. Several promising approaches, such as METHONTOLOGY [8] and ROMEO [25], have been developed to provide clear structural guidelines, in order to simplify the complexities and eliminate doubts in ontology design and evaluation. However these have not yet been deployed in the context of sustainability

[1]http://www.globalreporting.org/Home
[2]http://www.oecd.org/home
[3]http://www.un.org/esa

indicators. The current study showcases an effort to build a robust and reusable ontology for sustainability indicators, by incorporating these approaches into an integrated framework for design and evaluation.

In this paper, we present an *ontology for sustainability indicator sets (OSIS)*, for representing information about sustainability indicator systems. In Section 2, we first review the METHONTOLOGY ontology engineering and the ROMEO ontology evaluation approaches. Section 3 discusses the development and support activities given from METHONTOLOGY to design two candidates for OSIS. This is followed in Section 4 by the evaluation activity using ROMEO to validate two ontology models with a set of experiments. Section 5 presents the results and our findings from performing the experiments. Lastly, Section 6 concludes this work and discusses future work to pursue.

2. RELATED WORK

Domain-level ontologies have been developed in a wide range of disciplinary areas, with bioinformatics and the life sciences being particularly prominent examples. Relatively little attention has been devoted to establishing ontologies in the domain of sustainability science. Kumazawa et al. [16] demonstrate one such example, using a problem-based approach to structure a complex conceptual hierarchy in support of a general knowledge base. Similarly, Brilhante et al. [5] show a bottom-up approach to ontology design to support economic sustainability indicators.

Our work builds upon both examples. While we adopt, as orienting principles, two of the requirements listed by Kumazawa et al. [16] – interpretability and reusability – our ontology is designed to reflect as much as possible extant sustainability indicators, rather than our conceptualisation of the sustainability domain. In that respect, it follows the 'bottom-up' approach of Brilhante et al. [5]. However the overall goal of OSIS is to support *multiple* indicator systems as easily as possible. Hence our emphasis is upon the conceptual apparatus of such systems themselves, rather than the sustainability domain they represent. This results in OSIS having concepts such as "Indicator". To support this goal, we also use existing ontology engineering and evaluation approaches.

2.1 Ontology Engineering

Ontology engineering refers to a set of principles that are related to the development of an ontology for a specific domain. Janssen et al. [15] introduce an ontology engineering process that consists of *setup*, *design*, *approval* and *dissemination* phases. Various case studies have used aforementioned phases for their particular application. In 2000, the Gene Ontology (GO) Consortium [1] was established to produce a dynamic and controlled vocabulary in the context of the biomedical data that can be applied as knowledge of gene and protein roles. The GO[4] consists of three independent ontologies. Ryu and Choi [23] present an approach for term and taxonomy extraction based on information theory. Ferfeleder et al. [10] also introduce a semantic guidance to facilitate requirement categorisation process which is a challenging problem in the field of requirement engineering. Their semantic system provides a list of suggestions for engineers to define requirements effectively by using concepts, relations and axioms of a domain ontology.

Among various ontology engineering approaches, METHONTOLOGY is chosen for OSIS development due to its high rate of adoption among other domain ontologies. This approach is presented by Gómez-Pérez and her colleagues in the context of developing an ontology in the domain of chemicals [8]. METHONTOLOGY's framework enables the ontology engineers to construct the ontology at the knowledge level, and the design lifecycle is based on evolving prototypes consisting of three distinctive activities – namely **Management**, **Development** and **Support** – activities, which are performed either serially or in parallel. Management activities, such as *control* and *quality assurance*, are conducted at the start of the ontology development to identify the tasks to be performed, and the time and the resources required for their development activities shape the basis of the ontology. Ideally these phases – *specification, conceptualisation, formalisation, implementation* and *maintenance*– are performed through small incremental and iterative cycles. Finally, support activities which are carried out simultaneously with the development activities include: *knowledge acquisition, evaluation, documentation* and *configuration* [8]. We focus only upon the initial design and development of OSIS candidates – discussed in detail in Sections 3 and 4 – while some phases such as management, maintenance and documentation, although vital to the general creation of ontologies, are not discussed here.

2.2 Ontology Evaluation

Several approaches exist to evaluate ontologies, including: *Gold standard*, *Criteria-based* and *Task-based*. The first approach compares an ontology with a benchmark ontology. Maedche and Staab [18] propose a gold standard approach to empirically measure similarities between ontologies from different views such as lexical and conceptual aspects. Criteria-based approach evaluates the ontology based on the specific criteria such as *consistency*, *completeness*, *conciseness*, *expandability* and *sensitivity* [11]. Additionally, different researchers have proposed various ontology evaluation criteria [11, 12, 13]. The task-based approach evaluates an ontology based on the competency of the ontology in completing tasks, for example, whether the measured performance of that ontology within a specific application yields out the suitability of the ontology for that application.

One example of task-based ontologies is ROMEO (Requirements Oriented Methodology for Evaluating Ontologies) [25] which focuses on requirements as tasks. ROMEO is a compatible ontology evaluation technique which links generic requirements, such as "competency", "capability", "functionality" and "standardized" to evaluation measures through the development of some criteria "questions". This method is discussed in detail in Section 4.

3. DESIGNING WITH METHONTOLOGY

Here we step through the activities specified by METHONTOLOGY to support the development of two candidate ontologies for sustainability indicators. For clarity, we separate these activities into two processes, relating to the Pre-Design and Design phases of ontology development respectively.

[4]http://www.geneontology.org

Domain (Subject)	Property (Predicate)	Rang (Object)
Superclass: Indicator+	dc:title dc:type dc:description dc:periodOfTime dc:publisher osis:instance-of osis:hasCategory osis:hasReference osis:hasUnitofMeasurement osis:hasIssue osis:belongToIndicatorSystem	String String Superclass: Description String Superclass: IndicatorSet Superclass: Indicator Superclass: Category Superclass: Reference String SuperClass:Issue SuperClass:IndicatorSet
SuperClass: IndicatorSet+	dc:title dc:text osis:hasIndicator	String String Superclass: Indicator
Subclass: Issue	dc:title osis:isMeasuredByIndicator	String Superclass: Indicator
SuperClass: Description+	dc:title dc:text dc:format dc:date dc:publisher	String String String Date Superclass: IndicatorSet
SuperClass: Category+	dc:title dc:text osis:hasIndicator	String String Superclass: Indicator
SuperClass: Reference+	dc:title dc:text dc:isReferencedBy	String String Superclass: IndicatorSet
* : Abstract Class + : Concrete Class		

Table 1: OSIS-Design A – Generic Model

Domain (Subject)	Property (Predicate)	Rang (Object)
Superclass: Indicator*	dc:title dc:type dc:description dc:creator dc:periodOfTime osis:instance-Of osis:hasCategory osis:hasReference osis:hasUnitofMeasurement	String String Superclass: Description Superclass: Author String superclass: Indicator Superclass: Category Superclass: Reference String
Subclass: GRI_Indicator+	gri:hasCompilation gri:hasDefinition gri:hasDocumentation gri:hasRelevance gri:hasAspect	Subclass:GRI_Description Subclass:GRI_Description Subclass:GRI_Description Subclass:GRI_Description Subclass:GRI_Aspect
SubClass: OECD_Indicator+	oecd:hasDefinition oecd:hasInformation oecd:hasTheme	Subclass: OECD_Description Subclass: OECD_Description Subclass:OECD_Theme
SuperClass: Description*	dc:title dc:text dc:format dc:date osis:instance-Of	String String String Date superclass: Description
SubClass: GRI_Description+	gri:hasIndicator	Superclass: Indicator
SubClass: OECD_Description+	oecd:hasIndicator	Superclass: Indicator
SuperClass: Category*	dc:title dc:text osis:instance-Of osis:hasIndicator	String String superclass: Category Superclass: Indicator
SubClass: GRI_Aspect+	gri:hasIndicator	Superclass: Indicator
SubClass: OECD_Theme+	oecd:hasIndicator	Superclass: Indicator
SuperClass: Reference*	dc:title dc:text dc:isReferencedBy osis:instance-Of	String String Superclass: SustainabilitySet superclass: Reference
* : Abstract Class + : Concrete Class		

Table 2: OSIS-Design B – Specific Model

3.1 OSIS Pre-Design Process

While METHONTOLOGYdoes not mandate a specific order, prior to beginning the formal design of OSIS we undertook *Specification* and *Knowledgement Acquisition* activities – as discussed below – which these provide key inputs into that design.

3.1.1 Specification

Gómez-Pérez et al. [8] state the purpose of the Specification phase is to produce an Ontology Requirements Specification Document (ORSD) in natural language, using informal, semi-formal or formal representation of the ontology. The ORSD identifies the scope, purpose and requirements of the ontology. It also specifies the level of formality required for the ontology, depending on whether terms and their meanings of the specific domain need to be codified in natural or formal language. We developed a specification of the OSIS ontology with the following purpose:

- **Purpose:** The aim of the ontology is to represent knowledge about sustainability indicators in the context of a specific application. The ontology can be reused for reasoning, reapplying and querying indicators for integration purposes.

3.1.2 Knowledge Acquisition

To translate the specification into design, we also conducted a Knowledge Acquisition stage, which METHONTOLOGY [8] emphasises as pivotal to developing a suitable domain ontology. Firstly, we consulted with sustainability experts associated with our broader project, through a series of interviews and workshops. Secondly, we analysed a number of available sources of domain knowledge, including widely used indicator sets, to extract key domain concepts to include in the ontology. Two indicator sets, the GRI and the OECD, were used to inform the initial design. The third system, taken from the UN, we selected as a frame of reference to evaluate OSIS.

Each system reflects subtle yet distinctive features of how sustainability is conceptualised by their respective organisations. Hence using 2+1 frames of reference allows us to triangulate the key domain concepts to at least a first degree of approximation. From this activity we identified the following key concepts for sustainability indicators including: *Core and Additional Indicators, Category, Description, Reference, Sustainability Set, Issue, Target, Objective, unitOfMeasurement, Title, ID, Organisation.* These key terms form the basis of the OSIS design we outline next.

3.2 OSIS Design Process

Having specified requirements and developed a set of key terms, we then began to design OSIS. Here we focus on three main phases suggested by the METHONTOLOGY approach [8] discussed below. These stages move us from a generic and abstract model of the domain through to progressively more specific design. During conceptualisation, we decided upon the need to develop two distinct conceptual models, each reflecting one of the two main requirements identified during the Specification phase. These two models are then formalised and implemented in the next two phases, which result in the candidate ontology designs shown in Table 1 and Table 2.

3.2.1 Conceptualisation

Conceptualisation is the most important step in ontology design [8]. The outcome of this phase is a specification of the ontology components, including key *concepts, relations and instances*, which should reflect a set of terms produced in knowledge acquisition phase (see Section 3.1.2). A key task here is determining how specific indicator system data, taken from GRI and OECD, should be specified in relation to abstract concepts of "Organisation", "Indicator" and so on. Such relations ideally should reflect the requirements of the final ontology design.

We begin by developing a *taxonomy* of concepts, that is a hierarchical structure, representing concepts and the appropriate relationships between these concepts. A concept is an entity with a key role that can be inherited by other entities. In the taxonomy, we make no distinction between the ontological status of conceptual entities – whether for instance they ought to be represented by classes or individuals, or related by generic properties or specific sub-class relations.

We then represent these concepts as *classes* in the ontology. Following the taxonomic organisation, such concepts are distinguished between those which are broad and abstract (such as "Organisation", "Indicator"), which we model as superclasses of the ontology, and those which are specific and thus form the subclasses. Subclasses are inherited by superclasses, using *is-a* relationships. Additionally, classes may have properties that complement them. Where such properties express relationships between classes, such classes express a *has-a* relationship.

A key difficulty in translating such principles of ontology design is deciding whether relationships are of the *is-a* or *has-a* kind. Such decisions are influenced by different factors that also pertain to Object Oriented (OO) design in software engineering. In OO design, a class represents a set of *objects* – the terms object and instance are interchangeable – that share a common structure and a common *behavior* [3]. An object captures some well-defined behaviour from its class but it has a unique identity. The key point in designing an OO model is identifying the structure and behaviour of similar objects which construct the classes. Therefore, defining an entity as a class depends on whether its instances (objects) have common properties (behaviors).

In designing OSIS, this principle is reflected in the treatment of specific indicator classes, depending upon the emphasis. From the point of view of *reusability*, we see that system indicators can be included as instances of anonymous classes that extend "Indicator", and that are further specified by particular properties (e.g. "belongsToIndicatorSystem"). From the point of view of *explicitness*, we also see a usefulness in specifying particular system indicators as subclasses (e.g. "GRI_Indicator") of a generic indicator class (e.g. "Indicator"). The first view is more broad to cover sustainability indicators' key information with no reference to any particular organisations which is called ***Design A*** (shown in Table 1). The second view is more detailed to include direct references to specific indicator sets, which is called ***Design B*** (shown in Table 2). Both designs are discussed as follows:

1. ***Design A:***
   In this design, we sought to define broadly a suitable conceptual structure which reflects the generic key information of sustainability indicators. To a large extent, sustainability indicators are introduced to address issues of critical conditions in complex systems. In other words, indicators can provide solutions for such issues.

2. ***Design B:***
   In this design, our emphasis is on the organisations that develop sustainability indicators. Here we include as key conceptual constructs these organisations and their own indicator classifications. Therefore, a range of classes and relationships is specifically added for each sustainability indicator set.

3.2.2 Formalisation

This phase involves the transformation of the conceptual models defined previously into a formal representation as an ontology. This involves both the explicit representation of concepts and relations as classes, properties and individuals, and the development of *namespaces* for grouping related entities. Like *F-logic*, which is the basis of the ontology conceptualisation, namespaces are used to distinguish similar properties and relationships used in various ontologies from each other. The namespace declaration in ontology is similar to XML, where an alias is associated with the URI of a conceptual resource. We use the following namespaces in our designs:

- ***osis:*** refers to the URI[5] used to represent the most abstract and generic concepts and relations, such as *<osis:hasIndicator>* and *<osis:hasTarget>*.
- ***dc:*** refers to the Dublin Core metadata URI[6] to label common properties that pertain to most or all entities, for example *<dc:title>* for name entities and *<dc:type>* for type entities.
- ***gri:*** refers to the Global Reporting Initiative for the properties that are specifically related to GRI sustainability indicator sets, such as *<gri:hasAspect>* and *<gri:hasDocumentation>*.
- ***oecd:*** refers to the Organization for Economic Cooperation and Development for the properties that are specifically related to OECD sustainability indicator sets, for instance *<oecd:hasDefinition>* and *<oecd:hasTheme>*

The latter two namespaces are used only in Design B.

3.2.3 Implementation

To implement the conceptualisation, we first selected a suitable formal language to represent two OSIS design candidates. A knowledge representation language must have four essential features: "vocabulary", "syntax", "semantics" and "rules of inference" [17]. We have decided to represent the ontology in Resource Description Framework (RDF) [7] and Web Ontology Language (OWL) because of two reasons: 1) RDF/OWL ontologies are easily extensible by others and 2) RDF/OWL ontology data can be reasoned by computational agents and description logics using existing querying and visualisation tools, such as ontology editors, SPARQL libraries and third-party vocabularies.

A) Technical Implementation:
We also adopted particular technology tools to store and model the triples. PostgreSQL [8] was selected as a triple store

[5] http://www.cs.rmit.edu.au/knowledgebase/ontology/OSIS#
[6] http://purl.org/dc/elements/1.1/
[7] http://www.w3.org/RDF/
[8] http://www.postgresql.org/

framework, due to its compatibility with both relational (SQL) and semantic (RDF and OWL) languages. This simplified the loading of indicator sets, which were often expressed in loosely or unstructured forms, and therefore needed to be manipulated into formats such as CSV. We also used Protégé [9] to edit the two ontologies.

We then converted the two candidates into RDF/OWL form using Protégé. This involved converting each concept and relationship into equivalent semantic triple statements of *subject*, *object* and *predicate*. Each subject (e.g. a class or instance) is linked to an object (another class or instance) by a predicate (e.g. a "is-a" property). Subjects are considered as the *Domain* concepts of the property or relationship and objects are considered as the *Range*. Once specified, both ontologies were exported from Protégé into SDB2 and Postgres.

B) Metadata Document Management:
Once completed the technical implementation, we loaded both GRI and OECD indicator sets as instance data into both ontology candidates.

The GRI sustainability indicator set is presented in *eXtensible Business Reporting Language (XBRL)* [10]. XBRL is an XML-based language introduced to exchange business information. It uses the XML notation such as XML schema, XLink and XPath to express the semantic connections required in business reporting. The GRI organisation uses this language to define their sustainability metadata in a taxonomy that captures the individuals reporting concepts as well as the relationship between concepts and other semantic meanings in the original document. The details of our data transformation approach are given in Section 4.2.1.

4. EVALUATING WITH ROMEO

Having developed the two OSIS candidate ontologies and populated them with sustainability indicator data, we evaluated them with the four stages adapted from METHONTOLOGY.

4.1 OSIS Evaluation Process

In METHONTOLOGY, each ontology is evaluated with a collection of ontology frames of reference. Yu et al. [25] define a frame of reference F, $F = < F_c, F_i, F_r >$, where F_c is the set of concepts, F_i is a set of instances and F_r is a set of relationships $-$ which is the union between the set of relationships between concepts F_{cr} and the set of relationships instances F_{ir} $-$ in a frame of reference. Here we interpret "frame of reference" to be the sorts of knowledge sources solicited during the Knowledge Acquisition activity – namely, the indicator systems themselves. Since the OSIS candidates are designed to support two key requirements of intuitiveness and reusability, accordingly we have chosen three indicator sets to evaluate the candidates against. The first two are the sources used to construct the candidates, the GRI and the OECD indicator systems. We refer to this below as the 'seen' frame of reference. The third one is one of the other sources, the UN indicator system. We refer to this as the 'unseen' frame of reference (since neither candidate has any explicit entities drawn from it). To conduct the evaluation, we designed several experiments to test both candidates against the high-level requirements described in Section 3.1.1, and further application and end-user requirements elicited through discussion with project stakeholders.

4.1.1 Establishing the Ontology Role

According to Yu et al. [25], eliciting the roles of an ontology is important to understand how the ontology is used in the context of an application and it also helps to determine a set of appropriate ontology requirements. The role of OSIS is defined as follows:

- **Ontology Role:** Enhance effectiveness of query expansion module in suggesting indicators for query tasks.

4.1.2 Ontology Requirements

Ontology requirements reflect a specific competency or quality of the ontology that can be obtained from existing ontology requirements or application requirements. In the context of sustainability indicators, we define two ontology requirements based on the ontology role and purpose.

- **Ontology Requirement 1:** Does the ontology provide a precise and intuitive representation of the indicator systems it represents?

- **Ontology Requirement 2:** Does the ontology allow for other indicator systems to be easily incorporated using existing concepts, properties and relations?

4.1.3 Criteria Questions

The ROMEO approach stipulates that a set of questions is administered for each of the requirements. Such questions explore various aspects of a given requirement providing a deeper understanding of the ontology. In addition, criteria questions lead to appropriate measures which are critical in an ontology evaluation context. Yu et al. [25] propose a list of criteria questions for a variety of ontology requirements. We specify two questions for the OSIS ontology candidates, and ensure each question is answered with respect to both seen (GRI and OECD) and unseen (UN) frames of reference. The criteria questions are listed below:

1. Do the ontology components (concepts, instances and relationships) adequately cover the terms of the given domain?

2. Do the ontology components (concepts, instances and relationships) capture the terms of the given domain correctly?

The first question examines the *coverage* criteria and the second question determines the *correctness* feature of the ontology.

4.1.4 Measures

At the final stage, Yu et al. [25] suggest adopting a set of measures that are compatible with the ontology requirements which allow us to answer the criteria questions. The ontology evaluation literature has proposed various ontology criteria and measure that are summarised in a previous survey conducted by Brank et al. [4]. Of these, we adopt the *precision* measure presented by Guarino [14] to measure the correctness criterion, by determining the percentage of overlapping terms in an ontology O that overlaps with the set of terms from a frame of reference F (Equation 1). Additionally, *recall* [14] is used to measure the coverage criterion,

[9]http://protege.stanford.edu/
[10]http://www.fxasustainability.com.au/

Example of a GRI indicator	EN2 **Percentage of materials used that are recycled input materials.** 1. Relevance This Indicator seeks to identify the reporting organization's ability to use recycled input materials... 2. Compilation 2.1 Identify the total weight or volume of materials used as reported under EN1... 3. Definitions Recycled input materials: Materials that replace virgin materials that are purchased or obtained from internal or external sources... 4. Documentation Potential information sources include billing and accounting systems, the procurement or supply management department... 5. References OECD Working Group on Waste Prevention and Recycling.
Example of XBRL for the above indicator	<**label xlink:type**="resource" **xlink:label**="gri-core_EN02_lbl_en_terseLabel" **xlink:role**="http://www.xbrl.org/2003/role/terseLabel" **xml:lang**="en" **id**="gri-core_EN02_lbl_en_terseLabel">EN2</**label**> <**label xlink:type**="resource" **xlink:label** ="gri-core_EN02_lbl_en_label" **xlink:role**="http://www.xbrl.org/2003/role/label" **xml:lang**="en" **id**="gri-core_EN02_lbl_en_label">EN2</**label**> <**label xlink:type**="resource" **xlink:label**="gri-core_EN02_lbl_en_guidelineDefinition" **xlink:role**="http://www.globalreporting.org/2006/G3/guidelineDefinition" **xml:lang**="en" **id**="gri-core_EN02_lbl_en_guidelineDefinition"> Percentage of materials used that are recycled input materials. (Core)</**label**> <**label xlink:type**="resource" **xlink:label**="gri-core_EN02_lbl_en_protocolRelevance" **xlink:role**="http://www.globalreporting.org/2006/G3/protocolRelevance" **xml:lang**="en" **id**="gri-core_EN02_lbl_en_protocolRelevance"> This Indicator seeks to identify...</**label**> < **label xlink:type**= "resource" **xlink:label**="gri-core_EN02_lbl_en_protocolCompilation" **xlink:role**="http://www.globalreporting.org/2006/G3/protocolCompilation" **xml:lang**="en" **id**="gri-core_EN02_lbl_en_protocolCompilation"> 2.1 Identify the total weight or volume of materials used as reported under EN1. 2.2 Identify the total weight or volume of recycled input materials. If estimation is required ...</**label**> ...
RDF triples after applying SAX on XBRL file	<EN2> <http://www.w3.org/1999/02/22-rdf-syntax-ns#**type**> <GRI_Indicator>. <EN2> <http://purl.org/dc/elements/1.1/**description**> "Environment". <EN2> <http://purl.org/dc/elements/1.1/**format**> "pdf". <EN2> <http://purl.org/dc/elements/1.1/**isReferencedBy**> <EN2-Reference>. <EN2> <http://purl.org/dc/elements/1.1/**title**> "Percentage of materials used that are recycled input materials". <EN2> <http://www.cs.rmit.edu.au/knowlodgebase/ontology/OSIS/**hasUnitOfMeasurement**> "percentage". <EN2> <https://www.g1oba1roporting.org/**hasAspect**> <EN2-Aspect>. <EN2> <https://www.g1oba1roporting.org/**hasRelevance**> <EN2-Relevance>. <EN2> <https://www.g1oba1roporting.org/**hasCompilation**> <EN2-Compilation>. ... <EN2-Relevance> <http://purl.org/dc/elements/1.1/**text**> "This Indicator seeks to identify the reporting organization's ability to use recycled input materials...". <EN2-Compilation> <http://purl.org/dc/elements/1.1/**text**> "2.1 Identify the total weight or volume of materials used as reported under EN1...". ...

Table 3: An example of a GRI indicator and a snapshot of its XBRL and RFD representation

referring to the percentage of overlap between a set of terms from the ontology and the frame of reference (Equation 2).

$$precision(O, F) = \frac{|F \cap O|}{|O|} \quad (1)$$

$$recall(O, F) = \frac{|F \cap O|}{|F|} \quad (2)$$

These metrics are originally given from Information Retrieval literature – known as *retrieval performance evaluation* [2] – to determine the quality of the answer set generated from a query task. However they are also applied in ontology context. For example, Rodriguez and Egenhofer [22] use these metrics to measure the fraction of similar entity classes from different ontologies, and Euzenat [9] also uses such metrics for measuring ontology alignments.

The F-measure is the harmonic mean of precision and recall that provides a sense of adequacy and balance modelling of the domain being presented in the ontology. Yu et al. [25] apply the F-measure to ontology evaluation in the context of indicating appropriate coverage of concepts in the relevant frame of reference. They give Equation 3 as follows:

$$F\text{-}measure(O,F) = \frac{2}{\frac{1}{recall(O,F)} + \frac{1}{precision(O,F)}} \quad (3)$$

4.2 Experiments

In evaluating the OSIS, we perform six sets of experiments. We select indicators from the category of *Economy* from three frames of reference (the GRI, the OECD and the UN indicator sets). We then compare F-measures for both OSIS ontology candidates, Design A and Design B, against the frames of reference.

4.2.1 Preparing and Analysing Documents

To show these steps in practice, we include an example of a GRI indicator from the Environment category and its XBRL and RDF representation in Table 3. As described in section 3.2.3, the GRI sustainability indicator set is represented in XBRL. First, we phrased the required data from this document using SAX functions including: *start-document, start-element, end-element, character and end-document* by applying on *label* tags with conditioning the relevant attributes such as id, $xlink : type$ and $xlink : label$. Second, the ontologies are populated with indicator data with the use of Protégé interface, for instance some of the ontological properties as described in Table 2 and shown in bold font in Table 3 are *description, format, isReferencedBy*. Next, we export the RDF file – the final ontology document – from Protégé. An example of the generated file in *NTriples*[11] is shown in Table 3 that describes the triple statements consisting of subject, predicate and object which represents the key information of $EN2$ using relevant namespaces, concepts and appropriate relationships in Design B.

The final step of the experiment obtains the overlapping terms between the original frames of reference and the ontology documents. In order to produce consistent results, a pre-processing stopping algorithm and Porter's Stemmer technique [20] are applied to the resulting overlapping terms. We also use the algorithm presented by Broder [6], which determines the syntactic similarity between two documents. In our case, the first document is the textual representation of the ontology, and the second is the frame of reference. Each document is considered as a sequence of tokens, divided into the number of contiguous subsequences called *shingles*, of length n that is also known as *n-gram*. The algorithm compares the sets of *n-grams* from two documents and calculates their resemblance value.

[11] http://www.w3.org/2001/sw/RDFCore/ntriples/

Frame of Reference	$\|F\|$	OSIS Ontology	$\|O\|$	$\|F\| \cap \|O\|$	$Precision_{ave}$	$Recall_{ave}$	$F\text{-}Measure$
GRI−Frame	2560	*Design−A*	2280	1602	0.702	0.625	**0.661**
	2560	*Design−B*	2309	2090	0.905	0.816	**0.854**
OECD−Frame	986	*Design−A*	802	590	0.735	0.598	**0.659**
	986	*Design−B*	890	765	0.859	0.877	**0.867**
UN−Frame	500	*Design−A*	445	325	0.650	0.733	**0.682**
	500	*Design−B*	303	247	0.494	0.315	**0.307**

Table 4: Results for OSIS, GRI Frame, UN Frame, precision, recall and F-measure

5. RESULTS

In Table 4, we present results from experiments described in the previous section. We use the average F-measure along with associated metrics, the average recall and precision, to compare the two OSIS design candidates against seen (GRI and OECD used in Design B) and unseen (UN not used in Designs A and B) frames of reference.

The number of terms ($|F|$) between the three frames is different due to a number of reasons. For example, in a comparison with the GRI and the UN, while both frames distinguish between economic, environmental and social indicators, the GRI is directed largely towards corporate sustainability reporting, while the UN indicator set is aimed at measuring nation-level sustainability development. The GRI therefore includes more economic indicators, while the UN emphasises social indicators. The UN set also includes a fourth category of 'institutional' indicators, which inflates the overall indicator count. This consequently affects the number of terms in each ontology $|O|$ and overlapping terms $|F| \cap |O|$ for each set of experiments.

The graph in Figure 1 features the results for coverage using the F-measure. Comparing the results for the GRI and the OECD with the UN frames reveals similar coverage (approx. 65%) for Design A. By contrast, the F-measure shows significant difference between the two frames for Design B; the GRI-Frame and OECD-Frame have large proportions of coverage (on average 85%) whereas the UN-Frame's number declines significantly (30%).

These figures can be explained by comparing the design decisions for the two candidates. Design A presents a generic model for the ontology with no direct reference to any sustainability indicators. Actual indicators are assumed to instantiate, rather than inherit, from the *Indicator* class. We view this accordingly as a more terse and generic conceptualisation of the domain. By contrast, Design B presents a specific model, with more class references to particular sustainability indicators. This results in a higher F-measure against the seen frames of reference (GRI and OECD) – but because the specific wording of concepts maps directly to that frame, it performs more poorly against the unseen frame of reference (UN).

The contrasting results map intuitively to different requirements that can be said to underwrite the construction of the two ontology candidates. Where an ontology needs to be *precise* and *transparent* – to faithfully represent, at a conceptual level, the specific conceptualisation of given indicator system – the approach adopted by Design B is preferred, since it results in better F-measures where that system is used as a frame of reference. In contrast, Design A better supports cases where the requirements emphasise ontology *reuse* with minimal cost of extension or refactoring, since it performs better against unseen frames of reference. We also acknowledge there are cases where a compromise between these requirements might result in a hybrid of Designs A and B. Indeed, such an option (maximising F-measures against both seen and unseen frames of reference) might be preferred where cost and time constraints permit.

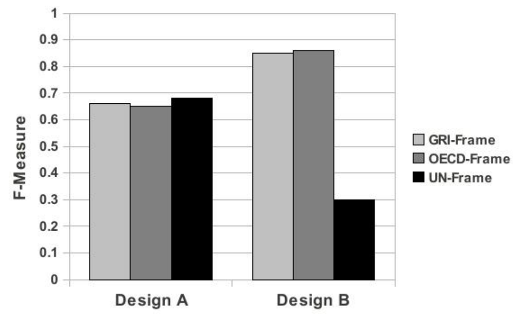

Figure 1: Results showing granularity using F-measure, GRI and OECD are used in constructing the ontology and UN is not used.

6. CONCLUSION AND FUTURE WORK

In this paper, we briefly introduce the field of sustainability indicator systems, and argue that ontologies are well-suited for representing such systems formally. We adopt METHONTOLOGY, a well-known ontology engineering methodology, to guide the development of two pilot ontology candidates for this domain. We then apply ROMEO to evaluate the candidates. The evaluation consisted of precision and recall to test the degree of coverage of indicators against two frames of reference. These metrics are also used in other ontology evaluation research [9, 14, 22, 25]. The first two frames of reference (GRI and OECD) were used to construct Design B; the third (UN) was only used in the evaluation.

The two candidates, A and B, differ largely in terms of abstraction. Design A applies an object-oriented style approach. Here, for example, the concept *indicator* is defined as a class, while specific instances of indicators are treated as individuals which instantiate properties and relationships of the *Indicator* class. By contrast, Design B treats each indicator instance as a class as well. Accordingly, they inherit rather than instantiate properties and relationships of the *Indicator* class. This produces a much larger ontology that maps directly to the specific frames of reference that it is derived from. Accordingly, Design B scores higher F-measure results against the *seen* frames of reference (the GRI and the OECD have been used in constructing this model). However, as our results show, Design A produces a better F-measure score against an *unseen* frame of reference, such as the UN that has not been used in informing this model.

We conclude that Design B is preferable where the domain requirements require a high degree of fidelity to seen frames of reference, while Design A offers greater reuse in contexts where unseen sets of indicators need to be added to the ontology in an ad hoc fashion. As sustainability indicators themselves continue to evolve, for this specific domain we argue Design A is preferable; though as consensus builds among reporting organisations, we also anticipate the possibility of blending both approaches in future. More generally, we show that both METHONTOLOGY and ROMEO can be productively used to guide the design and evaluation of domain-level ontologies, and that quantitative measures such as F-measures can be used to develop heuristics for preferring one ontology candidate to another, given a set of requirements and frames of reference.

We anticipate further work along several lines. Firstly, we think that ROMEO can be linked to METHONTOLOGY in a more systematic way, to guide ontology development from requirements through to evaluation and selection. Secondly, ROMEO itself can be extended through the sorts of quantitative procedures we apply here. In addition, a user-study is required to further evaluate the two design candidates with the real-world scenarios. In order to receive feedback from expert and non-expert users, one scenario can be applying Designs A and B on a knowledge base of a web-based application that reasons and queries indicators for various purposes. Finally, with reference to the field itself, further work can be undertaken to incorporate additional sustainability indicators systems, and to further refine the candidate OSIS ontologies presented here.

7. REFERENCES

[1] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, 25(1):25–29, 2000.

[2] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, 1999.

[3] G. Booch. *Object-oriented analysis and design with applications*. Benjamin-Cummings, 1994.

[4] J. Brank, M. Grobelnik, and D. Mladenic. A survey of ontology evaluation techniques. In *Conference on Data Mining and Data Warehouses (SiKDD)*, 2005.

[5] V. Brilhante, A. Ferreira, J. Marinho, and J. Pereira. Information integration through ontology and metadata for sustainability analysis. In *The International Environmental Modelling and Software Society (iEMSs) 3rd Biennial Meeting*, 2006.

[6] A. Broder. Identifying and filtering near-duplicate documents. In *Combinatorial Pattern Matching*, volume 1848 of *LNCS*, pages 1–10. Springer, 2000.

[7] B. Chandrasekaran and T. R. Johnson. Generic tasks and task structures: history, critique and new directions. In *Second generation expert systems*, pages 232–272. Springer, 1993.

[8] O. Corcho, M. Fernández-López, and A. Gómez-Pérez. Methodologies, tools and languages for building ontologies.where is their meeting point? *Data and Knowledge Engineering*, 46(1):41–64, 2003.

[9] J. Euzenat. Semantic precision and recall for ontology alignment evaluation. In *Proc. 20th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 348–353, 2007.

[10] S. Farfeleder, T. Moser, A. Krall, T. Stålhane, I. Omoronyia, and H. Zojer. Ontology-driven guidance for requirements elicitation. In *The Semanic Web: Research and Applications*, volume 6644 of *LNCS*, pages 212–226. Springer, 2011.

[11] A. Gómez-Pérez. Towards a framework to verify knowledge sharing technology. *Expert Systems With Applications*, 4(8):519–529, 2007.

[12] T. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.

[13] N. Guarino. Some ontological principles for designing upper level lexical resources. In *1st International Conference on Language Resources and Evaluation*, pages 527–534, 1998.

[14] N. Guarino. Towards a formal evaluation of ontology quality. *IEEE Intelligent Systems*, 19(4):74–81, 2004.

[15] S. Janssen, F. Ewert, and L. Hongtao. Defining assessment projects and scenarios for policy support: Use of ontology in integrated assessment and modelling, environmental modelling and software. *Special issue on simulation and modelling in the Asia-Pacific region (ASIMMOD)*, 24:1491–1500, 2009.

[16] T. Kumazawa, O. Saito, K. Kozaki, T. Matsui, and R. Mizoguchi. Toward knowledge structuring of sustainability science based on ontology engineering. *Sustainability Science*, 4(1):99–116, 2009.

[17] H. J. Levesque. Foundations of a functional approach to knowledge representation. *Artificial Intelligence*, 23(2):155–212, 1984.

[18] A. Maedche and S. Staab. Measuring similarity between ontologies. In *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*, volume 2473, pages 15–21. Springer, 2002.

[19] D. Maynard, W. Peters, and Y. Li. Metrics for evaluation of ontology-based information extraction. In *Workshop on Evaluation of Ontologies for the Web(EON), Scotland*, 2006.

[20] M. F. Porter. An algorithm for suffix stripping. In *Readings in information retrieval*, pages 313–316. Morgan Kaufmann Publishers Inc., 1997.

[21] D. Richards. The reuse of knowledge: a user-centred approach. *International Journal of Human-Computer Studies*, 52(3):553–579, 2000.

[22] M. Rodriguez and M. Egenhofer. Determining semantic similarity among entity classes from different ontologies. *Knowledge and Data Engineering, IEEE Transactions on*, 15(2):442–456, 2003.

[23] P. Ryu and K. Choi. Taxonomy learning using term specificity and similarity. *Workshop on Ontology Learning and Population*, pages 41–48, 2006.

[24] L. Steels. The componential framework and its role in reusability. In *Second generation expert systems*, pages 273–298. Springer-Verlag, 1993.

[25] J. Yu, J. A. Thom, and A. Tam. Requirements-oriented methodology for evaluating ontologies. *Information Systems*, 34(8):766–791, 2009.