Undefined 1 (2009) 1–5
IOS Press

Development and Evaluation of a Software System for Engaged Sustainability Reporting

Editor(s): Name Surname, University, Country
Solicited review(s): Name Surname, University, Country
Open review(s): Name Surname, University, Country

Lida Ghahremanlou [a], Liam Magee [b]. and James A. Thom [c]
[a] *IBM Research Australia, 204 Lygon Street, Carlton VIC 3053, Australia*
E-mail: lida.ghahremanlou@au1.ibm.com.au
[b] *University of Western Sydney, Locked Bag 1797 Penrith NSW 2751, Australia*
E-mail: l.magee@uws.edu.au
[c] *RMIT University, GPO Box 2476, Melbourne VIC 3001, Australia*
E-mail: james.thom@rmit.edu.au

Abstract. Measuring sustainability is a critical task for both public and private institutions. While developing a set of organisational indicators is challenging – there are competing definitions of, and measurements for, what sustainability means – a software system can simplify the identification of indicators from existing resources. However, the problem of the semantic heterogeneity of indicators, with shared, overlapping, or more specialised meanings, remains at the user interface level. We tackle this problem by developing software for a sustainability reporting framework – Circles of Sustainability – which provides two interface mechanisms for browsing indicator sets: a Tabular view and a Circular view. In developing such software, we introduce a generic design of an ontology for representing sustainability indicators. This ontology supports the presentation of indicators from multiple sources. We then evaluate the overall usefulness and ease of use of the presented software and the associated interface mechanisms by conducting a user study experiment. The quantitative and qualitative analysis in the user study indicates that the Circular view is the interface preferred by most participants to browse semantic heterogeneous indicators. Preferred aspects include the improved graphical display and visual experience, ease of navigation and visual communication, and providing short details about indicators by distinguishing their domains and sub-domains. The Tabular view, however, has advantages of simplicity and ease of keyword searching.

Keywords: sustainability indicators, user interface evaluation, ontology engineering

1. Introduction

Integration and reconciliation of data represented by heterogeneous resources is a pressing topic in different domains of science. It can be described as a problem of *semantic heterogeneity*, referring to various representations of the same or overlapping data. One type of semantic heterogeneity can be seen in databases when data duplicated across multiple resources needs to be integrated. Modern database management systems are unable to solve semantic heterogeneity problems that are distinguished in two categories [25].

Heterogeneity appears through different approaches to the physical layout of databases including configurations and network protocols [14]. From the point of view of logical structure of a database, such issues appear at the data and schema level; various conceptualisations and database schemas [11,19,32] are proposed in response.

The second another type of semantic heterogeneity issues can be seen in web documents, which require the integration of structured and semi-structured data resources [4]. In this form, not only semantic representation of data is an issue, but also the variation of data structures represented by different resources adds to the complexity. To address these, some solutions have been suggested in the development of methods, techniques and languages for web content [6,7,9].

More recently, the proliferation of data systems on the web has caused larger challenges for the integration of knowledge systems. For instance, Pisanelli et al. [43] discuss issues with regards to the integration and sharing of an extensive amount of information stored in various system repositories, in which each system has its own structure. Data integration in this context is related to *interoperability*, when the collaboration between systems and groups increase. A good example is the communication and exchange of information, which is a key requirement in business applications [4].

In a parallel argument, Magee [34] raises the concern of assessing the *commensurability* of knowledge systems by developing an implied theoretical approach accompanied by a heuristic and analytic framework. Magee [34] suggests using the concept of *ontological culture* in constructing formal knowledge systems by their communities (cultures) to represent vague kinds of entities and the development of *interlanguages* for translation between the micro-languages within knowledge systems to reduce their exponential complexity.

1.1. Heterogeneity in Sustainability Indicators

Increased attention to the sustainability and sustainable development [37,8] has encouraged the emergence of sustainability reporting. The term *sustainability reporting* refers to a general set of descriptive reporting practices, from “top-down” annual reporting against standardised indicator sets, to *ad hoc*, one-off or semi-periodic assessments against “bottom-up” and locally-grown measures [18]. Top-down approaches produce generalised sustainability reports with a global focus, while bottom-up approaches tend to be adopted by non-government organisations (NGOs), sub-national or municipal authorities, and community groups [1]. The similarity of both approaches is the use of *framework-based* techniques including a collection of measures, procedures, tools and principles that guide reporting practices.

Sustainability indicators are often introduced to address issues of critical conditions in complex natural and social systems. Indicators can help to provide solutions for such issues addressing three key objectives [30]:

1. Raising awareness and understanding of a system.
2. Informing decision-making.
3. Measuring progress toward established goals.

As these objectives suggest, sustainability indicators are commonly developed for the purpose of organisational governance, where they are often developed as simple taxonomies rather than fully fledged ontologies.

The issues of semantic heterogeneity of sustainability indicator sets appear at two levels: *data* and *interface*. At the data level, the issues concern representing multiple sustainability indicator sets, in which various organisations use unstructured or semi-structured data using different vocabularies to represent common concepts and relations of the sustainability domain. The second type of heterogeneity problem appears in browsing and navigating sustainability indicators at the user interface level. This is the second type of the focus of this paper.

Some examples of eneralised standard indicator sets are: the Global Reporting Initiative (GRI)[1], and the Organization for Economic Co-operation and Development (OECD)[2] and the United Nations Statistics Division (UN Social Indicators)[3]. The GRI reporting framework is designed to serve organisations' economic, environmental, and social sustainability performance of any size, sector or location. The main advantage of the GRI reporting framework is the practical considerations faced by a diverse range of organisations from small enterprises to those with extensive and geographically dispersed operations. The GRI reporting framework guideline contains general and sector-specific content agreed by a wide range of stakeholders around the world. The mission of the Organisation for the Economic Co-operation and Development (OECD) is to improve the economic and social well-being of people around the world. The OECD reporting on sustainability provides a forum in which government can work together to share experiences and seek solutions to common problems at the *country* level.

Semantic technology can be used to develop solutions for heterogeneity issues in the domain of reporting sustainability indicators. The main focus of this work is to address the research question of:

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

What mechanisms can be used to help end-users navigate and browse semantic heterogeneity of sustainability indicator sets?

We demonstrate the development of a software system we have termed Circles of Sustainability (CoS), following a method proposed by James [26] as an associated web-based software that embeds the methodology and model to construct a meaningful indicator-reporting framework. This software system offers two ways of assisting users to manage the semantic heterogeneity of sustainability indicators. In prototyping the Circles of Sustainability user interface, we suggest two methods of browsing and navigating sustainability indicators: a Tabular view and a Circular view, in which each interface is designed to assist end-users to browse aspects of heterogeneous indicator systems. We then design and conduct a user study to evaluate the overall usefulness of the CoS user interface and ease of use of the two browsing mechanisms.

The structure of the paper is as follows. We begin by reviewing the related work in Section 2. It is followed by outlining the CoS structure as a sustainability reporting framework in Section 3. Next, we describe the development process of the CoS user interface in Section 4. We then describe the steps for evaluating the CoS user-interface with a user study in Section 5. Section 6 discusses quantitative and qualitative results of the user study. Finally, Section 7 concludes the research findings.

2. Related Work

Related work includes text-based and visual interfaces, evaluation of user interfaces, and ontologies and taxonomies used for sustainability indicator sets. We discuss each of these in the sections that follow.

2.1. Mechanisms for Browsing Heterogeneity of Sustainability Indicators

The displaying methods are the Tabular and the Circular views which are text-based and visual interfaces respectively.

A review of interactive user interviews for browsing generic linked data are described by Alahmari et al. [2], and Dadzie and Rowe [13]. Text-base interfaces are considered easier for search entities because of more intuitive rendering and navigating features. Some examples include: Sig.ma [10], SView [12] and SWSE [23]. Visual interfaces are used individually or in combinations, using graphical structures such as images, maps, graphs and timelines to represent information. Some examples include: VisiNav [22] and DBpedia Mobile [3].

Our own evaluation incorporates suggestions for these general investigations of linked data navigation in the context of sustainability indicators. We compare the textual Tabular view and visual Circular view. The implementation details of these are given in Section 4.3.

2.2. Evaluating User Interfaces

The evaluation of user interfaces involves a process of *usability inspection*, which aims at identifying and addressing usability problems of an entire design. This topic has been increasingly applied as a method to evaluate *user interfaces* since the 1990s. Two broad views for evaluating user interfaces are: *assisting in design decisions* and *measuring the quality of use* [40]. The first perspective affects the designers' decision-making of possible alternatives (such as new user interfaces), and the second reports on the value of some measures for comparison purposes, which has resulted in the emergence a field of comparative evaluation approaches for user interfaces. In comparative evaluation, some studies[29,41] use multiple-criteria decision making to characterise the overall usability of an interactive system. Other studies[16,40] apply an integrated assessment of several interface characteristics to measure the usability.

Davis [15] studies the conceptual theories and empirical distinction of these variables from several diverse lines of research, such as self-efficacy theory, cost-benefit paradigm, adoption of innovation, evaluation of information reports and channel disposition model. The user behaviour is also examined from the point of view of vendor organisations. Some industrial instances are: IBM Watson[4], Xerox[5], Digital Equipment Corporation[6]. A common approach is A/B testing[7], in which, user perception is influential in decisions to use specific information technologies. The widespread use of A/B testing has led to "usability testing" becoming a standard phase of the product development cycle.

[4]http://www.ibm.com/smarterplanet/us/en/ibmwatson/
[5]http://www.xerox.com/
[6]http://www.digital.com/
[7]http://www.smashingmagazine.com/2010/06/24/the-ultimate-guide-to-a-b-testing/

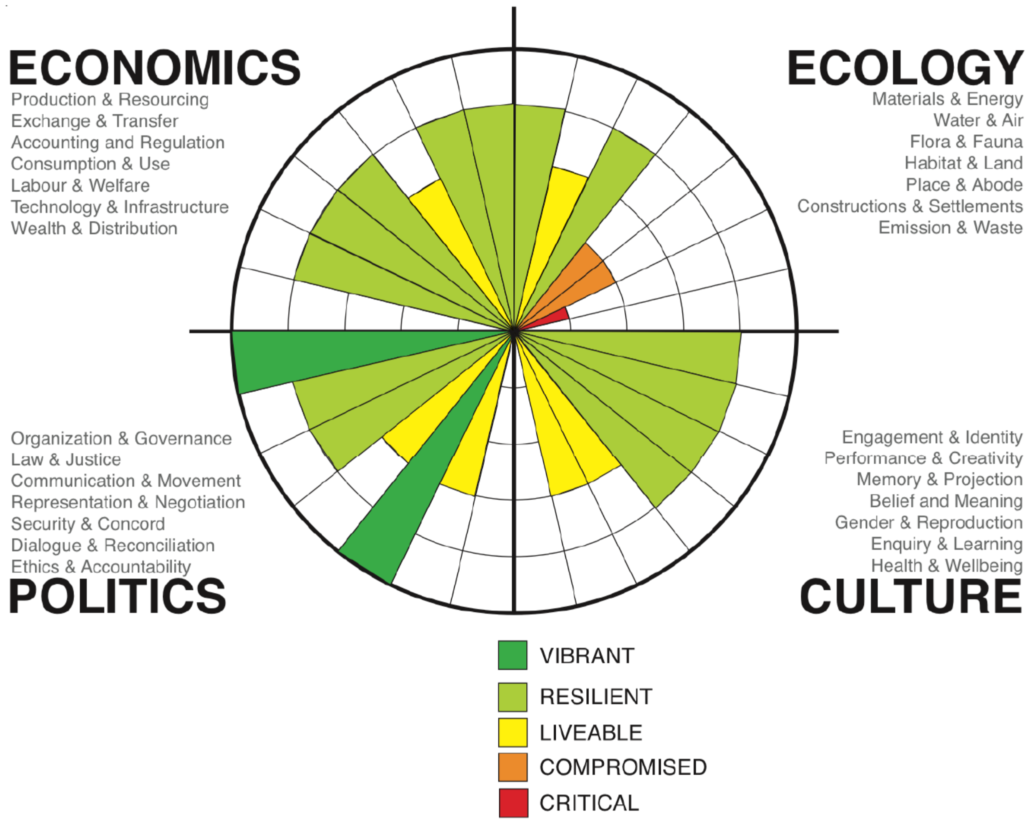

Fig. 1. Circles of Sustainability Framework (domains and subdomains) [35]

The International Organisation for Standardisation [17] defines usability as the extent "to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction". The common dimensions of usability, *effectiveness*, *efficiency*, and *satisfaction*, are used to elicit evaluative data from subjects about these qualities. A survey conducted by Hornbæk [24] on current practice of Human Computer Interaction research defines each dimension and lists several *usability measures* for each dimension. Effectiveness refers to accuracy and completeness of a tool which helps users to accomplish particular tasks. Some examples of effectiveness measures are: "Binary task completion", measuring the number or the percentage of the tasks that user completes successfully; and "Quality of the outcome", measuring the quality of users' understanding and their learning of information provided in the interface. Efficiency measures the resources expended in relation to achieving the goals by the user. In other words, software is efficient if it helps users complete their tasks with minimum waste, expense or effort. Most measures of efficiency are time-based as time is an important factor in assessing various tools. Satisfaction is defined as the positive attitudes of the user to the product. It also refers to the fulfilment of a specific desire or a goal such as, the measuring gratification and contentment of the users' experience when they accomplish particular goals [28]. Some examples of satisfaction measures include: "Preference", measuring satisfaction as the interface users prefer using, and "Ease of learning", measuring the amount of effort expended learning to use a system.

We use satisfaction dimension to define the variables in the questionnaire of the user study – given in Section 5.2 – because it is a broad measure of users' overall satisfaction or attitudes towards the interface or user experience. We focus on two specific measures of *perceived usefulness* and *perceived ease of use*, that are the fundamental concepts of user acceptance presented by Davis [15]. Perceived usefulness is the degree to which a person believes that using a particular system would enhance his or her job performance, which is extracted from the definition of the word *useful*, "capa-

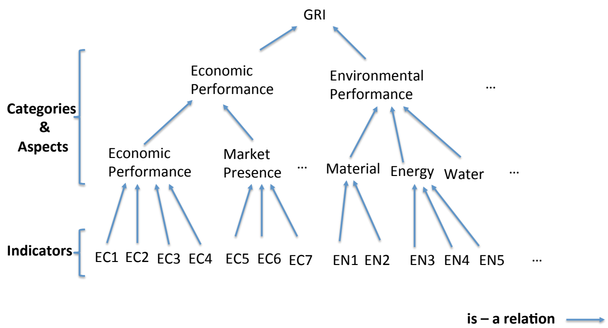

Fig. 2. A Snapshot of GRI Indicator Set

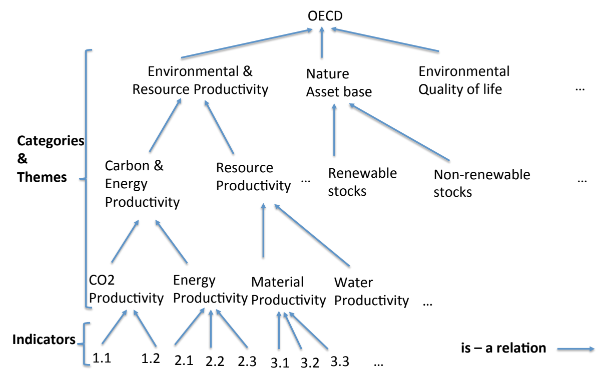

Fig. 3. A Snapshot of OECD Indicator Set

ble of being used advantageously". Perceived ease of use is the degree to which a person believes that using a particular system would be free of effort, which also follows from the definition of *ease* meaning "freedom from difficulty or great effort" [15].

A complete review of methods for evaluating information retrieval systems with users is presented by Kelly [28]. We choose a user study (survey) approach and follow the steps described by Kasunic [27] to evaluate the user interfaces of the CoS software to gain insight into users after working with the software. The details are given in Section 5.

2.3. *Ontologies and taxonomies used in Sustainability and Sustainability Indicator Sets*

There have been several attempts to develop domain and application ontologies in the context of sustainability and sustainability reporting indicator sets. Brilhante et al. [5] present an ontology for the domain of Indicators and Sustainable Development (ISDs)[8] with the emphasis on the economic dimension (ISD-Economics). This ontology is developed on the basis of the UN report on countries undertaking sustainable development programmes [30], which provides an appropriate vocabulary for the user interface with a hierarchical organisation of the domain. In addition, ISD-Economics operates on indicators using properties and relations defined in the ontology. Similarly, Madlberger et al. [33] develop an ontology for the domain of Corporate Sustainability Information Systems. This ontology fulfils three requirements: (i) Providing an intuitive representation of Corporate Sustainability concepts; (ii) Reflecting the hierarchical structure of economic indicators; and (iii) Generic mapping of sustainability heterogeneous resources – such as GRI-XBRL taxonomy – into the predefined concepts of the ontology. Several application ontologies have also been developed for specific tasks. Kumazawa et al. [31] outline an ontological approach to structure the concepts and relations within the field of sustainability science. Han and Stoffel [21] suggest an ontology for the integration of qualitative case studies in the domain of environmental sustainability research. Similarly, an ontological approach is presented by Pinheiro et al. [42] to link sustainability indicators with a social ontology [42] by relating and structuring environmental, social and economic issues in a way that measures overall sustainability.

Taxonomies refer as the centre of most conceptual models. According to Welty and Guarino [45], well-structured taxonomies are the key elements in reusing and integrating models and are particularly useful for human interpretation, whereas improperly structured taxonomies results in models which are confusing and difficult to reuse or integrate. In ontology development, a taxonomy is referred to a hierarchical structure of a set of concepts and subsumption relations between them. These concepts are then represented as *classes* in the ontology. Broad concepts compose *superclasses*, that are inherited by *subclasses* using `is-a` relation. Unlike ontology, in taxonomy there is no distinction between the conceptual entities and whether they ought to be represented by classes or individuals, or related to generic attributes or specific

[8]http://www.un.org/esa/sustdev/publications/indisdmg2001.pdf

relations. In other words, the additional relations between concepts are not represented in a taxonomy.

Taxonomies of sustainability reporting indicator sets include several standardised frameworks for corporate and other organisational reporting, notably the GRI and OECD. These indicator sets or systems represent sustainability indicators in hierarchical structures including categories and sub-categories. Snapshots of the GRI and OECD taxonomies are shown in Figures 2 and 3, compromising categories and sub-categories (or aspects and themes) and indicators at the bottom of the hierarchy. More specifically, the GRI sustainability indicator set is presented in *eXtensible Business Reporting Language (XBRL)*[9]. XBRL is an XML-based language introduced to exchange business information. It uses the XML notation such as XML schema, XLink[10] and XPath[11] to express the semantic connections required in business reporting. Some research has investigated how to transform such a taxonomy into an ontology. For example, Madlberger et al. [33] first used the GRI-XBRL taxonomy[12] as the basis to automatically generate the concepts of the Corporate Sustainability ontology.

3. Circles of Sustainability Framework

Circles of Sustainability framework is first introduced by Scerri and James [44]. Its methodology is later presented by Magee et al. [35]. In what follows we describe the domain model, the methodology, and associated web-based open-source software system for accompanying the methodology.

3.1. Circles of Sustainability Model and Methodology

The Circles of Sustainability model contains five main components:

1. Domains and subdomains.
2. Issue, normative goals and objectives.
3. Indicators, indicator sets and targets.
4. Network Relationships between issues, indicator sets and individual indicators.
5. Data collected against the indicators.

[9]http://www2.xbrl.org//au/
[10]http://www.w3.org/XML/Linking
[11]http://www.w3.org/TR/xpath/
[12]https://www.globalreporting.org/reporting/reporting-support/xbrl/Pages/default.aspx

The model uses four *domains* – Economics, Ecology, Politics and Culture – and associated *subdomains* illustrated in Figure 1. At the heart of these is the concept of an *issue*, that is in between the overall reporting context (such as an organisation, or key organisation project), the domain model, and the specific variables or indicators against which data is collected. Various issues can be related to the entire reporting scope (*general*), or be identified with some level of attention (*particular*). A general issue can correspond to a shared intention agreed within the system (*normative goal*). A consensus of the stakeholders may also express some particular desirable state of affairs (*objective*) for a subordinate issue. The objectives are expressed as a set of *indicators* and associated *targets* that are observable and measurable. The reporting framework requires collection of data to measure indicators related to issues of concern.

The purpose of the Circles of Sustainability methodology is to specify several rules for constructing the entities of the model. The key steps in the methodology are as follows:

1. Definition of general issue and associated normative goal
2. Determination of critical issues and relevant objectives
3. Selection of indicators and associated targets
4. Establishment of relationships between domains and subdomains, issue sets, issue and indicators
5. Reporting against chosen indicators
6. Developing and monitoring a response
7. Reviewing and adapting an indicator model

In this article, we use the first four steps to develop and evaluate the CoS software.

3.2. Circles of Sustainability Software

The CoS software incorporates the conceptual model outlined above in the form of a knowledge-base repository, including both locally-grown and internationally-standardised indicator sets. The software system makes use of an use-case-based architecture, with the aim of providing sophisticated reasoning support at various stages. For example, the software could recommend use of a particular indicator for a given issue, based on heuristics such as issue or objective name similarity, frequency of indicator use, provenance of the indicator – whether it is sourced from a standard indicator set such as the Global Reporting Initiative for example – and so on. A key affordance of this added reasoning support is the robust defensibility of relationships between issues and indicators; report compilers can readily retrieve and justify the particular rationales for measuring an issue with a chosen indicator.

This system similarly makes use of heuristics to correlate local, user-defined indicators to global standards like the Global Reporting Initiative. This feature exploits the semantic definition of indicators to infer possible concordances with global indicators. These concordances can be presented to users for review, and provide a pathway to the partial automation of standardised reporting via standards such as XML Business Reporting Language (XBRL)[13]. While no immediate panacea to the reconciliation of top-down and bottom-up reporting is available without considerable interpretative work by the participants, even in prototype form the system may help organisations struggling with building both relevant and compliant reporting frameworks.

An important aspect of the software system is its ability to be both goal directed, in attempting to guide the users towards a balanced set of indicators, and towards covering the different stages of the methodology, as well as reactive to user preferences. Consequently the software support system is not simply a menu driven interface that steps the user through fixed methodological steps. Rather it is able to respond to the user's goals, allowing them to enter indicators where they wish, but at the same time maintaining an overarching goal to guide and assist the user towards fulfillment of the various stages.

The front-end of the CoS software is the interface of the web-page. We apply the Responsive Web Design (RWD) approach [39] that aims at providing an optimal displaying mechanism for easy reading and navigation experience through the web pages with a minimum of resizing, panning, and scrolling across a wide range of devices from mobile phone screens to desktop monitors.

3.3. Applying Generic Ontology Design on CoS *Software*

In designing the ontology, our focus was presenting a generic model for sustainability data that covers broad key information of the domain and various indicator systems. This generic ontology model is the development of our earlier work presented by Ghahremanloo et al. [20] in 2012. The components identified in the ontology model correspond to the aforementioned conceptual entities of the CoS model – presented in Section 3.1. They are *domains, subdomains, issues, indicator sets* and *indicators*, in which "domains" and "subdomains" are replaced with Category and SubCategory due to various representations of the hierarchical structure of indicators in different sustainability reporting organisations. Description and Reference are also added for more clarification of the indicators.

[13]https://www.xbrl.org/

Table 1

Generic Ontology Design Applied on CoS Software

Concept (Domain)	Relation	Concept or Data Type (Range)
Superclass: Indicator	dc:title dc:type dc:description dc:periodOfTime dc:publisher osis:hasCategory osis:hasReference osis:hasUnitOfMeasurement osis:hasIssue osis:isBelongTo	String String Superclass: Description Date Superclass: IndicatorSet Superclass: Category Superclass: Reference String SuperClass:Issue SuperClass:IndicatorSet
SuperClass: IndicatorSet	dc:title dc:text osis:hasIndicator	String String Superclass: Indicator
Superclass: 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
SuprClass: Reference	dc:title dc:text dc:isReferencedBy	String String Superclass: IndicatorSet

From the viewpoint of *reusability*, we see that system indicators can be included as instances of `IndicatorSet` and an `Indicator` class is further linked by a particular relation (for example `belongsToIndicatorSet`). This view is more broader to cover sustainability indicators' key information with no reference to any particular organisations and is called *Generic Ontology for Sustainability Indicator Set*.

In this design [20], several sustainability reporting organisations are the values for `IndicatorSet` and an indicator belongs to one organisation. From the point of view of *reusability*, indicator systems can be included as instances of `IndicatorSet` and `Indicator` class which are further linked together by a particular relation (`belongsToIndicatorSet`). In addition, two namespaces are used in this model:

- **`osis:`** refers to the URI[14] used to represent the most abstract and generic concepts and relations, such as `<osis:hasIndicator>` and `<osis:hasDescription>`.
- **`dc:`** refers to the Dublin Core metadata URI[15] to label common properties that pertain to most or all entities, for example `<dc:title>` for name entities and `<dc:type>` for type entities.

Table 1 summaries the components of the generic ontology design that we apply on the CoS software.

4. Developing a Sustainability Reporting User Interface

In developing the CoS user interface, the following scenario is considered. A new user must first register. The user can then select or create a project associated with several critical issues. The user must then browse sustainability indicators from loaded indicator systems using either a Tabular view or a Circular view. Relevant indicators can then be added to critical issues. Finally, users need to add indicators and conduct an assessment of the project against the CoS.

4.1. Defining a Sample Project

For conducting the user study evaluation, we first create a sample project following the scenario previously described. The project considers **Tehran Air Pollution**, an issue that impacts all residents living in the city. Tehran is an extreme case of a community urban area that has several long-term sustainability problems that can result in short term disasters. Tehran, the capital city of Iran, is the largest urban area in that nation and had a population of nearly 8.5 million in June 2014[16]. According to Naddafi et al. [38], the city ranks as one of the largest cities in Western Asia and the 19th in the world. As with other large cities, Tehran is faced with serious air quality problems. 20% of the total energy of the country is consumed in Tehran. This results in the emission of pollutants of PM_{10}, SO_2, NO_2, HC, O_3 and CO_2, about 80-85% of which is produced by automobile sources of pollution [38].

4.2. Presenting a Series of Critical Issues

Based on preliminary research into this particular case, we pose the following issues concerning air pollution in Tehran. Using the terminology of our generic ontology described in Section 3.3, we specify these as *Critical Issues* and list them below. The description of the critical issues are given in Appendix A.

1. Geographical location of the city
2. The use of private cars
3. Poor public transport service
4. Low quality of the petrol made in Iran
5. Moving towards an industrialising city

4.3. Browsing Mechanisms of Indicator Sets

In pursuit of supporting the user navigation of semantic heterogeneity of sustainability indicator sets, two mechanisms of browsing indicators are used for translating the back-end into the front-end interface. We compare a text-based with a visual-based browsing approach. These are described further below.

The Tabular view is a text-based interface that presents information about indicators in a series of columns. This method provides a series of important attributes including: *IndicatorSet, Source, Category, Indicator, Title, Description* and *Subdomain(s)*. This view assists users to find information in a relative way and suggests an overall view of the indicator sets structure for those users who are less familiar with sustainability concepts. a keyword search area simplifies the process of finding relevant indicators.

By contrast, the Circular view is a visual interface that displays indicators in a circle divided into four *domains*, each of which is divided into seven sections (`Subdomains`). By clicking on a Domain the names of all *subdomains* appear at the left corner of the page and a list of related indicators and their details are illustrated at the right side of the page. In this view, reading subdomains is a good way to find relevant indicators, and is more suitable for those users who are more familiar with sustainability concepts.

Snapshots of both views are shown in Figure 4. In a comparison of the two views, the Tabular view is based on an existing software *component*, whereas the

[14] http://www.cs.rmit.edu.au/knowledgebase/ontology#
[15] http://purl.org/dc/elements/1.1/
[16] http://worldpopulationreview.com/world-cities/tehran-population/

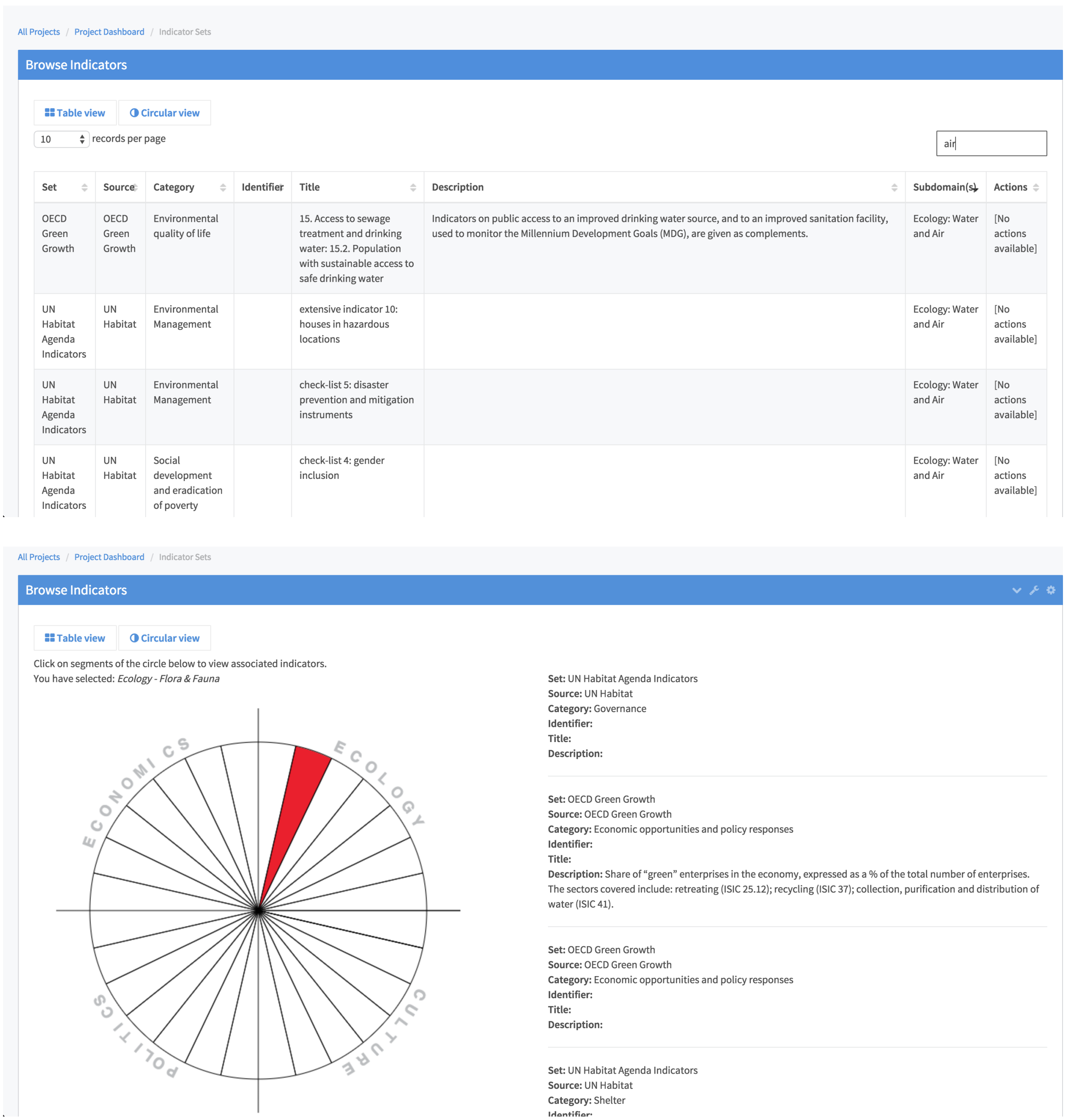

Fig. 4. Screenshots of Tabular View vs. Circular View

Circular view is a custom-developed *prototype* and features the CoS methodology shown in Figure 1, which recalls the concept of "Skeuomorphism" at user interfaces, in which a design is taken from the real world to recall the physical components[17]. The Circular view can be regarded as a conceptual representation of sustainability indicators based on the CoS methodology, whereas the Tabular view is a data conventional structural representation.

The GRI and OECD indicator sets were then loaded into the system. The software provided two options for assessing the sample project: first, by assessing added indicators against each of the CoS `subdomains`, or second, by entering values for each added indicator. Both the subdomain and indicator values form parts of an overall assessment of the project at a any given time.

5. Evaluating a Sustainability Reporting User Interface

To evaluate the user interface of the CoS software we conducted a user study, described below.

5.1. User Study Objectives

We have deployed two mechanisms (Tabular and Circular views) for browsing sustainability indicator sets in Section 4.3. Accordingly, a set of objectives is identified for the user study from perspectives of two types of participants, whom we discriminate into groups based on self-assessed familiarity with the sustainability domain. We term the first group *more-expert* and the second *less-expert*. We then determined two goals for the study. (i) The first goal aims to evaluate the usefulness of the CoS software overall. (ii) The second goal compares the two methods of browsing semantic heterogeneity of sustainability indicators that are designed in the CoS user-interface.

Reviewing the literature and considering two objectives of this research – overall usefulness of the software and comparing ease of use for two ways to browse the indicators – we decided on the *satisfaction* dimension and two specific usability measures: *perceived usefulness* and *perceived ease of use* as described in Section 2.2.

Based on the above objectives and two usability measures, three high-level variables are identified: *Expertise*, *Perceived usefulness of the software* and *Perceived ease of use* of two browsing mechanisms. These lead us to define the descriptive variables.

[17]http://www.techopedia.com/definition/28955/skeuomorphism

5.2. Descriptive Variables

Following the high-level variables, twelve descriptive variables are specified, which form the design of the questions and basis for analysing them.

Expertise

1. EX1: Prior knowledge of sustainability concepts, theories, approaches and methods.
2. EX2: Prior knowledge of indicator systems (such as the GRI, OECD and UN).
3. EX3: Prior knowledge of sustainability reporting frameworks and specifically the CoS methodology (for example project, critical issues, project assessment).

Overall Usefulness of the Software

4. OU1: Satisfaction level of using software for accomplishing tasks more quickly.
5. OU2: Satisfaction level of using software in sustainability assessment for improving users' performance, productivity and effectiveness.
6. OU3: Satisfaction level of easier use of software in sustainability assessment.
7. OU4: Satisfaction level of using the software in everyday sustainability assessment and reporting practices.

Ease of Use for Semantic Heterogeneity of browsing indicators

8. SH1: Satisfaction level to learn operating with both views more easily.
9. SH2: Satisfaction level to get both views to do what the user wants to do.
10. SH3: Satisfaction level of interaction with both views from clarity, comprehension and flexibility perspectives.
11. SH4: Satisfaction level in becoming skillful in using both views.
12. SH5: Satisfaction level to ease of use of both views.

5.3. Hypotheses

Prior to conducting the user study and collecting the results, we also defined several hypotheses by identifying the relations between high-level variables: expertise, overall usefulness of the software and semantic heterogeneity of browsing indicators, which also

Table 2
Mapping between Variables and Questions

High-Level Variables	Usability Measures	Descriptive Variables	Questions	Types of Questions
Expertise	**Background Knowledge**	EX1	Q1	**4-point Likert scale**
		EX2	Q2,Q3,Q4	
		EX3	Q5	
Overall Usefulness	**Perceived Usefulness**	OU1	Q6	**5-point Likert scale**
		OU2	Q7,Q8,Q9	
		OU3	Q10	
		OU4	Q11	
Semantic Heterogeneity	**Perceived Ease of Use**	SH1	Q13,Q19	**5-point Likert scale**
		SH2	Q14,Q20	
		SH3	Q15,Q16,Q20,Q21	
		SH4	Q17,Q23	
		SH5	Q18,Q24	

reflect two research objectives. The hypotheses are adapted from a study by McGrenere et al. [36] as follows.

Participant Hypothesis: Participants will probably have prior knowledge of sustainability domain and the CoS software. However, there will be more "less-expert" users than "expert" users who have limited knowledge about reporting organisations.

Usage Hypothesis: Most participants will find the CoS software as a helpful tool to accomplish their ability in sustainability assessment.

Good Idea Hypothesis: Two ways of browsing indicators will be easily understood and will be considered a good idea. It will be easy to switch between interfaces.

Satisfaction Hypothesis: Most participants will be satisfied with the CoS software overall.

Navigation Hypothesis: Most participants will benefit from general guidelines provided in the interface and progressively will feel that they are better able to perform the tasks.

Learnability Hypothesis: Using the Tabular view will provide more opportunities to learn the CoS software and its associated methodology.

Functionality Hypothesis: Using Circular view will facilitate most features of the CoS user interface.

Using these hypotheses enable us to analyse the data collected from the participants after conducting the experiment.

5.4. Questionnaire

After identifying the research objectives, variables and hypothesises of the user study, a set of questions are then developed for measuring the variables. In the current user study, we design a questionnaire that contains several closed and open questions. This allows us to collect quantitative information for particular subjects as well as gaining the insight into participants' responses. The subject of the questions associated with variables are given in Section 5.2. Each variable may correspond to one or more than one question(s). The questionnaire is divided into the three sections. The first contains 5 closed questions with a scale of 1-4 and assesses the background knowledge of the participants about sustainability concepts and reporting organisations. Participants are asked to answer this section before using the software and answering the other two sections. The second section of the questionnaire contains 6 questions in a 5-point Likert Scale and assesses the overall usability of the software. The third section contains two subsections for measuring "perceived ease of use" for the Tabular and Circular views respectively. In each subsection, 6 questions in the 5-point Likert Scale measure the users' experience of using each view from different aspects of usability, namely: accomplishing tasks quickly, improving performance, increasing productivity, enhancing effectiveness, performing easy assessment and usefulness.

Table 2 summaries the mapping between variables and questions. The types of the questions are also specified in the last row of the table. In addition, two open questions (Q12, Q25) and an overall feedback statement with the software are also given for collecting the qualitative responses. The time required to complete the experiment and answer the questionnaire is estimated approximately 20-30 minutes.

5.5. Characteristics of Target Audiences

After a series of discussions with our social scientist experts of the broader project about relevant user groups, we selected the participants who had prior knowledge about concepts and theories concerning the sustainability domain, although the levels of their background knowledge were varied. Other factors such as gender and age were not considered in the selection criteria. In recruiting participants, we approached individuals using the "snowball" approach. Most participants were invited through recommendations by people associated with the UN Global Compact Cities Programme[18]. These users were considered to be broadly representative of future users of the software, including members of universities, governments, community-based organisations and companies. On average our sample was, however, more aware about issues of sustainability than we would expect of most users. This may skew the results favorably, but may also cause users to be more critical of defects in the software.

5.5.1. Sample Size of Participants

Following the above approach, the number of potential participants reached to 38, although, 26 user submitted their responses. However, among them, only 20 user answered the three sections of the questionnaire.

Table 4
Descriptive Statistics of Six Questions - Perceived Usefulness

	N	Mean	Std. Deviation	Minimum	Maximum
Q6	20	1.800	.767	1.00	4.00
Q7	20	1.400	.502	1.00	2.00
Q8	20	1.900	.718	1.00	3.00
Q9	20	1.650	.587	1.00	3.00
Q10	20	1.600	.820	1.00	4.00
Q11	20	1.700	.864	1.00	4.00

[18]http://citiesprogramme.com/

5.6. Statistical Tests

The Spearman's Rank Correlation Coefficient (Spearman's rho) test is chosen for analysing the results of the second section of the questionnaire, measuring overall usefulness of the software. It is a non-parametric test and the significant difference indicates whether there is a correlation between the two variables by ranging from -1.00 to $+1.00$. This coefficient indicates the degree that low or high scores on one variable tend to go with low or high scores on another variable, where 1 represents a total positive correlation, 0 is no correlation and -1 is a total negative correlation. We used this test to identify strong correlations at significant difference of .05 between different aspects of usefulness of the CoS software.

The Wilcoxon Signed-ranked test is chosen to analysis the results for the third section of the questionnaire. It is also a non-parametric test and is used when the data cannot be assumed to be normally distributed. This test compares two sets of scores from the same participants investigating any changes from one time point to another, or when individuals are subjected to more than one condition. We used this test at significant difference of .05 to understand the preference of participants in working with two ways of browsing indicator sets (Tabular view versus Circular view).

5.7. Tasks

In developing the user study, participants were expected to accomplish the tasks below which are based on the scenario given in the development of the CoS user interface in Section 4.

1. Adding relevant indicators to issues of the sample project using both user interfaces.
2. Assessing the sample project and added indicators.

In order to guide the participants in performing these tasks, we also provided a set of instructions for using the CoS software and accomplishing the tasks within *Script A* and *Script B*. The scripts have similar steps with the only difference being in choosing which interface of browsing indicators to use at the beginning of the experiment. Script A asks users to choose the Tabular view first, while script B asks them to choose the Circular view first. Given the number of the participants, there is an even distribution of both scripts between them to eliminate the chance that the use of one view will dominate the responses (that is 13 users

Table 3
Spearman rho Correlations for Six Questions - Perceived Usefulness

		Q6	Q7	Q8	Q9	Q10	Q11
Q6	Correlation Coefficient	1.000	.656**	.515*	.486*	.715**	.745**
	Sig. (2-tailed)	–	.002	.020	.030	.000	.000
	N	20	20	20	20	20	20
Q7	Correlation Coefficient	.656**	1.000	.694**	.685**	.895**	.747**
	Sig. (2-tailed)	.002	–	.001	.001	.000	.000
	N	20	20	20	20	20	20
Q8	Correlation Coefficient	.515*	.694**	1.000	.657**	.549*	.477*
	Sig. (2-tailed)	.020	.001	–	.002	.012	.034
	N	20	20	20	20	20	20
Q9	Correlation Coefficient	.486*	.685**	.657**	1.000	.710**	.369
	Sig. (2-tailed)	.030	.001	.002	–	.000	.109
	N	20	20	20	20	20	20
Q10	Correlation Coefficient	.715**	.895**	.549*	.710**	1.000	.705**
	Sig. (2-tailed)	.000	.000	.012	.000	–	.001
	N	20	20	20	20	20	20
Q11	Correlation Coefficient	.745**	.747**	.477*	.369	.705**	1.000
	Sig. (2-tailed)	.000	.000	.034	.109	.001	–
	N	20	20	20	20	20	20

*Correlation is significant at the 0.01 level (2-tailed)
**Correlation is significant at the 0.05 level (2-tailed)

Table 5
Descriptive Statistics of Twelve Questions - Perceieved Ease of Use

	N	Mean	Std. Deviation	Minimum	Maximum
Q13	19	2.47	1.35	1.00	5.00
Q14	19	2.63	1.06	1.00	5.00
Q15	19	2.74	1.19	1.00	5.00
Q16	19	2.90	1.15	1.00	5.00
Q17	19	2.26	1.15	1.00	5.00
Q18	19	2.32	1.21	1.00	5.00
Q19	18	1.72	.83	1.00	4.00
Q20	19	1.80	.79	1.00	3.00
Q21	19	1.79	.63	1.00	3.00
Q22	19	1.79	.85	1.00	4.00
Q23	19	1.68	.88	1.00	4.00
Q24	19	2.32	1.20	1.00	5.00

received Script A and the other 13 received Script B). By alternating using two views, the participants' responses to the third sections of the questionnaire – related to comparing two views – are not biased to only one way of browsing indicators.

Next, we piloted the scripts and questionnaire with five users who had general knowledge of sustainability indicator sets and were partially familiar with the CoS software. The scripts were improved over some iterations by addressing the comments and feedback from pilot users.

We then conducted the user study and collected the results over a 10-day period in March 2014. The participants were contacted via individual emails and received an electronic copy of the scripts (A or B) and a link to the survey of the questionnaire.

6. Results and Discussion

In this section, we present quantitative and qualitative data collected from the user study experiment described in previous section. The subsequent analysis of the results are also provided. Quantitative data contains the participants responses to 23 closed questions from three sections of the questionnaire, while the qualitative data contains participants' comments and feedback to two open questions.

6.1. Quantitative Results

Here, we present the data collected from quantitative results. Focusing on the last two sections of the questionnaire, collected data is categorised into two distinct sub-categories as follows.

6.1.1. Overall Usefulness

The second section of the questionnaire measured perceived usefulness of the CoS software. This section had six questions with five response types: “Extremely Likely”, “Slightly Likely”, “Neither”, “Slightly Unlikely” and “Extremely Unlikely”, which ranked from 1 to 5 respectively. Lower values represents higher satisfactions of the participants with regards to the overall usability of the software. As discussed in Section 5.2, the questions’ topics correspond to the usability variables adapted from the study by McGrenere et al. [36]. We conducted the Spearman’s rho test to identify significant correlations between the usability variables and the questions are based on them.

Tables 4 and 3 illustrate the descriptive and correlation results in a tabular form. The sample size, shown by N, is 20. The results indicate that Q7 shows strong correlations with other five questions. In particular, the highest correlation value is between (Q7,Q10) of 0.895, which is significant at the confidence level of .01. The second highest correlation values are seen between pairs of (Q6,Q11) and (Q7,Q11) of 0.745 and 0.747 respectively, which are also significant at the confidence level of .01. In contrast, the weakest correlation exists between pairs of (Q9,Q11) and (Q6,Q9) with the values of 0.369 and 0.486 respectively.

These figures support some of our hypothesises given in Section 5.3 but rejects others. Q7 concerns whether participants’ performance is improved using the CoS software and it has strong correlations with all other questions, which focus on individual aspects of the usability of the software. This supports the satisfaction hypothesis that says most participants will be satisfied with the CoS software overall. Furthermore, Q11 suggests using the CoS software is a useful tool for sustainability assessment and has strong correlations with two other questions, which break down the usability aspects into two sub-tasks namely: Accomplishing tasks more quickly (Q6), Improving performance (Q7). This partially supports of the usage hypothesis that the CoS software will help the participants to accomplish the sustainability assessment. However, the weak correlations between pair (Q9,Q11) and pair (Q6,Q9) reject the hypothesis that using CoS will enhance the effectiveness of sustainability assessment. The reason could be the “effectiveness of the software” was not clear for all participants, who ranked it very low.

Table 6
Signed Wilcoxon Test - Sigfinicant Difference and Z-value - Perceived Ease of Use

		N	Mean Rank	Sum of Ranks
Q13-Q19	Negative Ranks	7^a	7.50	52.50
	Positive Ranks	4^b	3.38	13.50
	Ties	7^c	–	–
	Total	18	–	–
Q14-Q20	Negative Ranks	10^d	6.90	69.00
	Positive Ranks	2^e	4.50	9.00
	Ties	7^f	–	–
	Total	19	–	–
Q15-Q21	Negative Ranks	11^g	6.41	70.50
	Positive Ranks	1^h	7.50	7.50
	Ties	7^i	–	–
	Total	19	–	–
Q16-Q22	Negative Ranks	13^j	7.35	95.50
	Positive Ranks	1^k	9.50	9.50
	Ties	5^l	–	–
	Total	19	–	–
Q17-Q23	Negative Ranks	8^m	5.19	41.50
	Positive Ranks	1^n	3.50	3.50
	Ties	10^o	–	–
	Total	19	–	–
Q18-Q24	Negative Ranks	2^p	2.50	5.00
	Positive Ranks	2^q	2.50	5.00
	Ties	15^r	–	–
	Total	19	–	–

a. Q19<Q13, b. Q19>Q13, c. Q19=Q13
d. Q20<Q14, e. Q20>Q14, f. Q20=Q14
g. Q21<Q15, h. Q21>Q15, i. Q21=Q15
j. Q22<Q16, k. Q22>Q16, l. Q22=Q16
m. Q23<Q17, n. Q23>Q17, o. Q23=Q17
p. Q24<Q18, q. Q24>Q18, r. Q24=Q18

6.1.2. Ease of Use of Browsing Indicator Sets

The third section of the questionnaire aimed at comparing users’ satisfaction with the perceived ease of use of two ways of browsing indicators: Tabular view versus Circular view. Twelve 5-point Likert scale questions are paired for each view using the variables discussed in Section 5.2. The questions in this section also have the same type of the responses from previous section. For analysing the results, we used the Signed Wilcoxon test for the paired responses, for example Q13 (Tabular view) versus Q19 (Circular view).

Table 7
Signed Wilcoxon Test - Perceived Ease of Use

	Q13-Q19	Q14-Q20	Q15-Q21	Q16-Q22	Q17-Q23	Q18-Q24
Z	-1.749^b	-2.434^b	-2.516^b	-2.745^b	-2.326^b	.000^c
Asymp. Sig. (2-tailed)	.080	.015	.012	.006	.020	1.000

a. Wilcoxon Signed Ranks Test
b. Based on Positive Ranks
c. The Sum of Negative Ranks Equals the Sum of Positive Ranks

Tables 5 and 6 feature descriptive and rank results of the Wilcoxon test. The results suggest that questions related to the Circular view rank lower than Tabular view, in which, lower values in responses indicate greater satisfaction with a view. As a result, four pairs out of six questions – (Q14,Q20), (Q15,Q21), (Q16,Q22) and (Q17,Q23) shown in Table 7 – ranked positively at significant difference of .05. This indicates that users were more satisfied working with the CoS features using the Circular view rather than the Tabular view. Those features in which the Circular view is more preferred are: (i) Easy to understand what to do, (ii) Interaction is clearer and understandable, (iii) It is more flexible to interact with and (iv) It is easy to become skilful in its use.

Further analyses of the results indicates several findings. First, the Circular view is the preferred interface for dealing with the issue of semantic heterogeneity when browsing indicators by most participants. This confirms the functionality hypothesis, given in Section 5.3. Second, the learnability hypothesis of the Tabular view is not supported because the Wilcoxon signed test did not find any significant difference between the scores of the two views for the question of "Learning to operate with which view is more easier?". These findings lead us to a conclusion that the question, "Which view assists users more to use CoS software?" does not have an absolute answer. It is a relative answer that depends on which features of the CoS software are considered. Although, the Circular view is preferable for most reporting procedures, the Tabular view is also ranked to be a better alternative for some of the tasks. Subsequent analyses of qualitative participants' feedback clarify these points.

Q12. Can you imagine using the software in everyday sustainability assessment and reporting practices? What could be done or added to make it more useful?

The variety of responses to this question, lead us to divide them into two categories: Software strengths and suggestions to improve and software limitations, which are briefly summarized in Table 8.

Given the highlighted key points in Table 8 from participants' responses, it can be seen that overall the software is considered to be a useful tool for reporting practices and assessments, although there are conflicting ideas about the simplicity of using software. While some expert users benefit from the present scenario – defining a project and issues, adding indicators to issues and assessing the project against circles of sustainability – less expert users find this scenario confusing and complex. In addition, users with little knowledge about indicator sets seem confused about the adequateness of information provided in the interface. Furthermore, unclear navigation steps and less intuitive guidance are considered to be other shortcomings of the interface. This rejects the navigation hypothesis given in Section 5.3. Finally, users detected several technical issues concerning the current implementation.

6.2. Qualitative Results

Two open questions at the end of second and third sections of the questionnaire aim at understanding participants' insight. Here we highlight the key findings as follows.

Q25. The software ultimately aims to help an organisation to select indicators for measuring a series of critical issues that a stakeholder group has identified. In addressing this goal, which view is preferable?

This question investigates the insight of users' experiences of working with the two ways of browsing indicator sets at the end of third section of the questionnaire. In Table 9, nine features are listed for each view extracted from users' responses. While it is not appropriate to compare individual responses, some spe-

Table 8
Q12: Software Strengths and Limitations

Software Strengths and Improvement Suggestions	Software Limitations
1. The software can be useful in comparing different datasets with a sustainability reporting framework. 2. It can be improved by a guidance of methodology steps to assist the navigation. 3. The simplicity with which the user is able to navigate the site makes this an ideal tool for everyday use. 4. The buttons and text clearly guide the user through each step. 5. While the instructions are useful, it would not take a user long to learn the software. 6. It allows the user to correct and update information. 7. The software has the potential to become an important component of urban development efforts. 8. The CoS software would definitely assist and support the assessment and reporting practices and process as it gives an immediate, tangible and readable results through the assessments. 9. It would be very useful if software could generate assessment "tools". 10. A searchable "tag" system would make it easier to find relevant indicators for an assessment. 11. A search function should be added to the Circular view. 12. Have a structured framework against which a user can perform each assessment. 13. Some interface design could be applied if it is expected to apply to more than a well-trained user. 14. Assessing indicators is clearer than adding issues and the charting tool looks really useful and user-friendly. 15. It would be magnificent and save a lot of time if the system was fully populated with dozens of relevant indicators in each subdomain. 16. The process may be facilitated and make it more user-friendly through the dashboard selection. 17. The circles inside the process would be better having the subdomains named directly against them or coming up with spatial contiguity with the cursor.	1. The software is a little cumbersome to navigate and uncomfortable to use. 2. The guidance through the steps is not intuitive, which means practice and time is required to understand the steps. 3. User interface is not intuitive for untrained users. 4. The first task in the experiment (that is adding issues) is a bit confusing. 5. It is unclear what sort of assessment values (second task) to enter in the final screen: there is a great deal of variation between the indicators. 6. The method of selecting the indicators is still unwieldy. Indicator sets are difficult to quantify and requires more description. 7. A large part of the problem is that there are not enough relevant indicators in the system. 8. Several technical issues are also detected by users as follows. (a) Having a delete button at the entry point can be risky. (b) The automatic process of subdomain selecting by keywords is inappropriate. For example, some of the subdomains would be automatically chosen when the user writes the terms of the issue. (c) While a sub-domain is selected if another sub-domain is also chosen, selection changes but the selection of the former sub-domain (red part) remains fixed. (d) The system does not seem to save consistently. (e) The hover-over action that changes the sub-domain should be removed so that makes selecting a sub-domain. (f) The "Add a new issue" button needs to be separated off from the "Add a new indicator". (g) Unless a page is accessed through the "Add existing indicator" path, no actions are available from either Tabular view or Circular view.

cific conclusions can be gained after reviewing the responses from both views.

The Tabular view is preferred for its simplicity in presenting indicators in a table module, having a search functionality, and being less visually oriented. However, this view fails at complex search queries and displays too much text at a time, which is not visually pleasing. The Circular view is preferred by most users. It provides several advantages including: comparative visual experience of assessment process, better graphical display, being easier to navigate and communicate visually and providing short details about indicators with distinguishing their domains and sub-domains. One downside with the Circular view is that indicators must have been previously assigned to at least one of the Circles of Sustainability.

7. Conclusion and Future Work

The overall objective of the present research has been to tackle heterogeneity issues of data and visualisation of sustainability indicator sets. In this paper, we addressed two different mechanisms, informed by recent studies of linked data user interfaces. Our results suggest that both mechanisms can help to resolve the heterogeneity problems of sustainability indicators at the user interface level. Through the development of the CoS software, we applied the generic ontology design to model the indicator systems and suggested two mechanisms of browsing indicator sets. We conducted a user study with a questionnaire to analyse the participants' satisfaction with using the CoS software and user interface for two usability measures of per-

Table 9
Q25: Users' Insight of Working with two Views

Tabular view	**Circular** view
1. The Tabular view is easier to browse and compare among indicator sets. 2. The Tabular view is better for specific searching. 3. The Tabular view is less visually oriented and more complex to use. Its exactness makes it ideal for reporting specific values, while simultaneously challenging to understand during the first use. 4. For complicated models with numerous indicators, the Tabular view would be more efficient and usable. 5. The Tabular view is more preferred because there are many text related to each indicator and would be useful to view them all together at one. 6. The Tabular view as it represents the environment in a more holistic and realistic manner. 7. In Tabular view it would be helpful if indicators could be 'refined' and sorted based on criteria such as 'tags', keywords, themes etc. 8. In the Tabular view browsing is slow, and a keyword search did not throw up the kind of indicators. 9. There was too much text involved with the Tabular view, which made it difficult to obtain an overall picture of the different indicators.	1. It provides a comparative visual experience of the assessment process. 2. The Circular view is better for a graphic display which then could be use for reporting more easily, quicker and more pleasant way. 3. It is visual and communicates to the user more easily and it is visually easier to navigate. 4. It is easier to see several different indicators simultaneously. This is important for comparison of indicators as some are very similar. 5. The Circular view appears to assist general indicator browsing and give short details. Also indicators can be viewed according to the individual frameworks e.g. OECD. 6. The Circular view is far easier to use because it only requires the user's intuition to figure out. Each piece of the circle can easily be identified and manipulated. 7. It would be useful to have the option to select additional areas of information per each section of the circle by simply clicking with the mouse on a single segment of the circle. 8. The Circular view may be more appealing for users with a language barrier. 9. The Circular view works well as long as the indicator can easily be categories against the four parameters.

ceived usefulness and perceived ease of use. The questionnaire was divided into three sections: background knowledge (5 closed questions), overall usefulness of the software (6 closed and 1 open question) and ease of use of browsing sustainability indicator sets (12 closed and 1 open questions). Accordingly, we analysed the quantitative and qualitative results following the three perspectives.

First, the results of the background knowledge suggest that most participants had some familiarity with sustainability domain and the CoS methodology. However, only a few were self-assessed experts who knew about sustainability indicators sets. Second, the subsequent analyses of quantitative results of the second section of questionnaire established the satisfaction hypothesis; that is, most participants were satisfied using the software. Another conclusion is the proof of usage hypothesis: the CoS framework assisted users to accomplish tasks more quickly, increased their productivity and enhanced the effectiveness of sustainability assessment with the higher level of satisfaction. Further, a review of users' feedback suggests that the CoS software is generally considered to be a useful tool for reporting practices and assessments, although there are conflicting ideas about the simplicity of using software from more-expert and less-expert users' points of view concerning the simplicity of using the software. In addition, inadequate information about sustainability indicators, unclear navigation steps and less intuitive guidance are seen to be other shortcomings. This conclusion therefore rejects the navigation hypothesis, which is not surprising for a prototype system that aims to support a complex set of features.

The results analysis of the third section of the questionnaire indicates several key findings. (i) The functionality hypothesis is accepted through the Circular view, which is the preferred interface to browse semantic heterogeneity of indicators by most participants. (ii) Learnability hypothesis about the Tabular view is rejected. (iii) The answer to the question, "Which view assists users more to use CoS software?" does not have an absolute answer and it depends on the functions of the CoS software. Although the Circular view is preferred for most reporting procedures, the Tabular view is also considered to be a better alternative for some features. This suggests that the time and effort spent developing an alternative view and navigation device needs to be balanced by the profile of the users and the tasks they are looking to perform. For comparing indicators, the Tabular view is preferred. For exploring indicator types, and assigning values to them, the Circular view is preferred.

For the future direction of the research, a combined user interface of visual and textual methods for browsing sustainability indicator sets would greatly improve the data representation issues. In addition, a sustainability reporting framework requires a concrete guidance or a tour to train less-expert users and simplify the complex concepts of sustainability domain for them. These sorts of additions would help further simplify the task of developing an appropriate and relevant framework for reporting progress on sustainability objectives.

Acknowledgement

This research was a part of a linkage project on *Accounting for Sustainability: Developing an Integrated Approach for Sustainability Assessments*, supported by the Australian Research Council (ARC) which was as part of the development of the UN Global Compact Cities Programme methodology aims to develop a "semantic" approach – called Circles of Sustainability – using open source software and systems, for supporting sustainability reporting within organisations. There are six different partners associated with the project number of LP0990509: Angusta[19], Cambridge[20], City of Melbourne[21], Microsoft[22], Fuji Xerox[23] and Cities Program[24].

Appendix

A. Description of Critical Issues for Tehran Air Pollution project [38]

1. **Geographical location of the city:** Tehran's average altitude is 1,890 metres above mean sea level. The city is located in valleys and surrounded on the north, northwest, east and southeast by high to medium high of mountain ranges of up to 3,801 meters height. These ranges block the flow of the humid wind to the city and prevent the polluted air from being carried away from the city. Thus, during winter, the lack of wind and the cold air causes the polluted air to be trapped within the city [38].
2. **The use of private cars:** Citizens are not encouraged to use public transport. Naddafi et al. [38] report that the city has a capacity for 700,000 registered cars are on its streets on a daily basis. Cars account for 70% to 80% of the city's air pollution.
3. **Poor public transport service:** Tehran Metro carries more than 2 million passengers per day[25]. Other types of public transport, such as bus lines, are not fully developed. Furthermore, taxies are very expensive. However, this type of transport is still the most used type which also has a direct effect on the air pollution problem.
4. **Low quality of the petrol made in Iran:** As sanctions on imports have forced the country to turn to low-quality fuel, all vehicles in Tehran use the poor quality petrol. The government has tried to replace the fossil fuel consumption with the gas option by encouraging drivers to convert their cars to using this alternative. However, the conversion process is costly and most drivers do not find this option economic.
5. **Moving towards an industrialising city:** Tehran, as Iran's capital and one of the mega-cities in the world, is moving towards industrialisation to strengthen its economy. The actions to support this aspect of the city have polluted Tehran, both directly and indirectly. On the other hand, the severity of the air pollution at certain times of the year has forced the government to shut down the city for a few days by asking people to stay at their homes. For example, in November 2013, kindergartens, elementary schools and universities were closed for three days. This action has considerable drawbacks on the economy of Tehran and the country as a whole.

References

[1] Annika Agger. Involving citizens in sustainable development: Evidence of new forms of participation in the Danish Agenda 21 schemes. *Local Environment*, 15(6):541–552, 2010.

[19] http://www.angusta.com.au/
[20] http://cambridgecollege.com.au/
[21] http://www.melbourne.vic.gov.au/Pages/default.aspx
[22] http://www.microsoft.com/en-au/default.aspx
[23] http://www.fujixerox.com.au/
[24] http://citiesprogramme.com/

[25] http://in.reuters.com/article/2011/03/05/idINIndia-55349520110305

[2] Fahad Alahmari, James A Thom, Liam Magee, and Wilson Wong. Evaluating semantic browsers for consuming linked data. In *Proceedings of the Twenty-Third Australasian Database Conference-Volume 124*, pages 89–98. Australian Computer Society, Inc., 2012.

[3] Christian Becker and Christian Bizer. Exploring the geospatial semantic web with dbpedia mobile. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(4):278–286, 2009.

[4] Sonia Bergamaschi, Silvana Castano, Maurizio Vincini, and Domenico Beneventano. Semantic integration of heterogeneous information sources. *Data & Knowledge Engineering*, 36(3):215–249, 2001.

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

[6] Peter Buneman. Semistructured data. In *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, pages 117–121. ACM, 1997.

[7] Peter Buneman, Susan Davidson, Gerd Hillebrand, and Dan Suciu. A query language and optimization techniques for unstructured data. *ACM SIGMOD Record*, 25(2):505–516, 1996.

[8] Ian Burton. Report on reports: Our common future: The world commission on environment and development. *Environment: Science and Policy for Sustainable Development*, 29(5):25–29, 1987.

[9] Diego Calvanese, Giuseppe De Giacomo, and Maurizio Lenzerini. What can knowledge representation do for semi-structured data? In *AAAI/IAAI*, pages 205–210, 1998.

[10] Stéphane Campinas, Diego Ceccarelli, Thomas E Perry, Renaud Delbru, Krisztian Balog, and Giovanni Tummarello. The sindice-2011 dataset for entity-oriented search in the web of data. In *1st international workshop on entity-oriented search (EOS)*, pages 26–32, 2011.

[11] Michael J Carey, Laura M Haas, Peter M Schwarz, Manish Arya, William F Cody, Ronald Fagin, Myron Flickner, Allen W Luniewski, Wayne Niblack, Dragutin Petkovic, J Thomas, J. H. Williams, and E. L. Wimmers. Towards heterogeneous multimedia information systems: The Garlic Approach. In *Research Issues in Data Engineering, 1995: Distributed Object Management, Proceedings. RIDE-DOM'95. Fifth International Workshop on*, pages 124–131. IEEE, 1995.

[12] G. Cheng, S. Gong, X. Li, Q. Liu, Zheng L., W. Jiang, J.and Hu, and Yu Qu. Sview: Sharing views for browsing linked data. In *Online proceedings on the Semantic Web Challenge website*, pages 26–32, 2013.

[13] Aba-Sah Dadzie and Matthew Rowe. Approaches to visualising linked data: A survey. *Semantic Web*, 2(2):89–124, 2011.

[14] Chris J Date and Hugh Darwen. *A Guide to the SQL Standard*, volume 3. Addison-Wesley New York, 1987.

[15] Fred D Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340, 1989.

[16] Elif Kılıç Delice and Zülal Güngör. The usability analysis with heuristic evaluation and analytic hierarchy process. *International Journal of Industrial Ergonomics*, 39(6):934–939, 2009.

[17] International Organization for Standardisation. *ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part II: Guidance on Usability*. 1998.

[18] Evan DG Fraser, Andrew J Dougill, Warren E Mabee, Mark Reed, and Patrick McAlpine. Bottom up and top down: Analysis of participatory processes for sustainability indicator identification as a pathway to community empowerment and sustainable environmental management. *Journal of environmental management*, 78(2):114–127, 2006.

[19] Michael R Genesereth, Arthur M Keller, and Oliver M Duschka. Infomaster: An information integration system. In *ACM SIGMOD Record*, volume 26, pages 539–542. ACM, 1997.

[20] Lida Ghahremanloo, James A Thom, and Liam Magee. An ontology derived from heterogeneous sustainability indicator set documents. In *Proceedings of the Seventeenth Australasian Document Computing Symposium*, pages 72–79. ACM, 2012.

[21] Dong Han and Kilian Stoffel. Ontology based qualitative case studies for sustainability research. In *Proceedings of the AI for an Intelligent Planet*, page 6. ACM, 2011.

[22] Andreas Harth. Visinav: A system for visual search and navigation on web data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 8(4):348–354, 2010.

[23] Aidan Hogan, Andreas Harth, Jürgen Umbrich, Sheila Kinsella, Axel Polleres, and Stefan Decker. Searching and browsing linked data with swse: the semantic web search engine. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(4):365–401, 2011.

[24] Kasper Hornbæk. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2):79–102, 2006.

[25] Richard Hull. Managing semantic heterogeneity in databases: a theoretical prospective. In *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, pages 51–61. ACM, 1997.

[26] Paul James. *Urban sustainability in theory and practice: circles of sustainability*. Routledge, 2014.

[27] Mark Kasunic. Designing an effective survey. Technical Report CMU/SEI-2005-HB-004, DTIC Document, University of Pittsburgh, Software Engineering Department, 2005.

[28] Diane Kelly. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1—2):1–224, 2009.

[29] Pekka J Korhonen and Jukka Laakso. A visual interactive method for solving the multiple criteria problem. *European Journal of Operational Research*, 24(2):277–287, 1986.

[30] OJ Kuik and AJ Gilbert. Indicators of sustainable development. In JCJM Van Den Bergh, editor, *Handbook of environmental and resource economics*, pages 722–730. Edward Elgar Publishing Ltd, 2002.

[31] Terukazu Kumazawa, Osamu Saito, Kouji Kozaki, Takanori Matsui, and Riichiro Mizoguchi. Toward knowledge structuring of sustainability science based on ontology engineering. *Sustainability Science*, 4(1):99–116, 2009.

[32] Alon Levy, Anand Rajaraman, and Joann Ordille. Querying heterogeneous information sources using source descriptions. Stanford InfoLab, 1996.

[33] Lisa Madlberger, Andreas Thöni, Peter Wetz, Alexander Schatten, and A Min Tjoa. Ontology-based data integration for corporate sustainability information systems. In *Proceedings of International Conference on Information Integration and Web-based Applications & Services*, page 353. ACM, 2013.

[34] Liam Magee. A framework for assessing commensurability of semantic web ontologies. *Electronic Journal of Knowledge*

Management, 8(1):91–102, 2010.

[35] Liam Magee, Andy Scerri, Paul James, James A Thom, Lin Padgham, Sarah Hickmott, Hepu Deng, and Felicity Cahill. Reframing social sustainability reporting: towards an engaged approach. *Environment, Development and Sustainability*, 15 (1):225–243, 2013.

[36] Joanna McGrenere, Ronald M Baecker, and Kellogg S Booth. An evaluation of a multiple interface design solution for bloated software. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 164–170. ACM, 2002.

[37] Donella H Meadows, Edward Goldsmith, and Paul Meadow. Limits to growth: 30 year update(3rd ed.). White River Junction: Chelsa Green, 2004.

[38] Kazem Naddafi, Mohammad Sadegh Hassanvand, Masud Yunesian, Fatemeh Momeniha, Ramin Nabizadeh, Sasan Faridi, and Akbar Gholampour. Health impact assessment of air pollution in megacity of Tehran, Iran. *Iranian Journal of Environmental Health Science & Engineering*, 9(1):1–7, 2012.

[39] Kailashkumar V Natda. Responsive web design. *Eduvantage*, 1(1), 2013.

[40] Kyung S Park and Chee Hwan Lim. A structured methodology for comparative evaluation of user interface designs using usability criteria and measures. *International journal of industrial ergonomics*, 23(5):379–389, 1999.

[41] José MC Pereira and Lucien Duckstein. A multiple criteria decision-making approach to gis-based land suitability evaluation. *International Journal of Geographical Information Science*, 7(5):407–424, 1993.

[42] Wallace Anacleto Pinheiro, Ricardo Barros, Jano Moreira De Souza, Geraldo Bonorino Xexeo, Julia Strauch, Patricia Barros, and Melissa Campos. Adaptative methodology of sustainability indicators management by ontology. *International Journal of Global Environmental Issues*, 9(4):338–355, 2009.

[43] Domenico M Pisanelli, Aldo Gangemi, and Geri Steve. Ontologies and information systems: the marriage of the century? *New Trends in Software Methodologies, Tools and Techniques*, pages 125–133, 2002.

[44] Andy Scerri and Paul James. Communities of citizens and 'indicators' of sustainability. *Community Development Journal*, 45(2):219–236, 2010.

[45] Christopher Welty and Nicola Guarino. Supporting ontological analysis of taxonomic relationships. *Data & Knowledge Engineering*, 39(1):51–74, 2001.