A Flexible Approach for User Evaluation of Biomedical Ontologies

Gilbert Maiga and Ddembe Williams

There has been an emergence of various ontologies describing data from either clinical or biological domains. Associated with this has been the development of integration systems, attempting to combine such biomedical ontologies using various strategies to overcome issues of scope and differing levels of granularity. However, such ontology based integration systems still find little use in distributed computing environments. This is attributed to both the lack of knowledge about user needs for such systems and the absence of a general framework to assess their relative suitability for specific applications. This paper aims to bridge this gap by proposing such a reference framework. The framework draws on existing information systems and ontology evaluation approaches in relating user objectives to ontology characteristics. The paper suggests that such a framework bridges the gap by emphasizing the dynamics of a biomedical environment. The framework therefore includes feedbacks from the evaluation process to the user characteristics of the integrated systems. This reference framework was validated in a study using structured interviews and a survey. The results indicate that the framework is sufficiently flexible for evaluating ontology based biomedical integrated systems, taking into account the conflicting needs of different users interested in accessing complex libraries of biomedical data.

1. Introduction

Biomedical ontology based integrated systems (BOIS) bring together disparate sources of heterogeneous and rapidly changing biological and clinical data [Kohler et al. 2003; Kumar et al. 2006; Yuggyung et al. 2006; Rey-Perez et al. 2006; Sioutos et al. 2006]. Ontologies represent domain knowledge in a generic way. They provide a commonly agreed understanding of a domain, which may be reused, shared, and operationalized across applications and groups. They are used to create reusable models for integrating biomedical data in order to achieve interoperability between such sources. However, the widespread use of BOIS in distributed computing environments remains hampered by: 1) the lack of a general framework to assess their suitability for specific applications; 2) the lack of knowledge about user needs for such BOIS since ontologies are subjective knowledge artifacts in terms of time, place and cultural environment, as reflected in their design [Alani and Brewster 2006]. This underlines the difficulty in articulating specific properties to use in ranking ontologies since selection can depend on personal preferences of user requirements [ibid].
Ontology integration builds new knowledge structures by assembling, extending, specializing and adapting existing ontologies. This enables purpose-built ontologies to emerge [Pinto and Martins 2000]. The challenge for evaluating ontologies has become an important research and development endeavor for assessing and selecting the appropriate biomedical integration ontology for a given system, especially in a distributed computing environment like the semantic web. Comparing the effectiveness of integrated ontology based systems remains challenging due to lack of standard frameworks for evaluating them, given that they differ in function, and in their expected input and outputs [Natalya and Musen 2002]. Existing frameworks evaluate ontologies largely on the basis of the taxonomical structure. These are therefore not suited for assessing the functional and processual objects present in biomedical ontologies. Furthermore, there is a lack of knowledge about the metrics and properties [scope, inputs, processes, outputs and structure] users require when judging the general quality or suitability of an integrated ontology [Alani and Brewster 2006; Kalfoglou and Schorlmer 2003; Lambrix and Tan 2006]. As a result, the effective reuse and wide adoption of ontology-based biomedical integration systems by industry is not yet realized [Gangemi et al. 2005; Kalfoglou and Hu 2006]. The EON [2006] workshop helped to bring into focus this lack of standard evaluation approaches and metrics for determining quality in the increasing number of ontologies made available for the semantic web.

In this study, a mixed method research strategy combining quantitative and qualitative deductive approaches was applied. A deductive descriptive approach helped to identify general properties and ontological theories that support knowledge integration in biological and clinical integration systems. The properties that users require in an ontology for biomedical integration systems were investigated using structured interviews and a survey. This helped to determine the scope, inputs, processes, outputs and evaluation criteria important to users of an integrated biomedical system. The results were used to derive a flexible framework for user centered assessment of systems that integrate clinical and biological ontologies. The framework is seen as a dynamic system, with feedbacks from the evaluation process guiding improvement of the integration systems.

The rest of this paper is organized as follows. Section 2 discusses work related to information systems evaluation and biomedicine. Section 3 explores existing approaches to ontology evaluation, multicriteria evaluations and systems theory as key underpinning concepts for deriving the framework. Conclusions from the study are drawn in section 4.

2. Biomedical and Information Systems Evaluation

2.1 Integrating Biomedical Information with Ontologies

Ontology-based attempts to bridge the gap between clinical and biological information in order to achieve interoperability include work on ONTOFUSION...
[Rey-Perez et al., 2006], SEMEDA [Kohler et al. 2003], ASLER [Yugyung et al., 2006], colon carcinoma [Kumar et al. 2006], and cancer [Sioutos et al., 2006].

While the ONTOFUSION tool provides semantic level integration of genomic and clinical databases [Rey-Perez et al. 2006], it does not bridge across levels of granularity for biomedical data. Using principles of the basic formal ontology (BFO), Kumar et al. [2006] describe a framework for integrating medical and biological information in order to draw inferences across levels of granularity using the Gene Ontology. Yugyung et al. [2006] describe a method for medical ontology integration using an incremental approach of semantic enrichment, refinement and integration (ALSER) that depends on measures of similarity between ontology models. A terminology and description logic based framework for integrating molecular and clinical cancer-related information given by Sioutos et al. [2006]. It however leaves integrating with external sources largely unresolved. Despite these efforts, integrating biomedical data objects across structure, function and processes using ontologies remains challenging with lack of a single unifying approach (frame of reference) to ontology integration, against which users can assess an ontology based integration system.

Evaluation of ontology integration systems is a criterion based technical judgment guiding the construction process and any refinement steps for both the integrated and resulting ontologies [Pinto and Martins 2000]. Ontologies may be assessed by user ratings and reviews, and for general ontological properties. Evaluation is done during design, development and prior to use [Kalfoglou and Hu 2006]. Evaluation is important if ontologies are to be widely adopted for use in distributed computing environments [ibid]. The lack of a unifying evaluation framework for integrated systems remains an obstacle for ontology reuse and may hinder their adoption by industry and the wider web community [Alani and Brewster 2006].

Little work exists on evaluating biomedical ontology integration systems (BOIS). To evaluate both the integrated and resulting ontologies, Pinto and Martins [2000] recommend both technical and user assessment of the candidate ontologies by domain experts and ontologists respectively. These authors also recommend selection of candidate ontologies using strict (hard) and desirable (soft) requirements. Soft requirements provide flexibility, as they can be adapted to integration processes that take into account particular features during the choice of one ontology.

Evaluation of the resultant ontology can therefore be done according to criteria used for any ontology to meet assessment criteria of completeness, conciseness, consistency, expandability, robustness [Gomez-Perez and Pazos 1995], and clarity, coherence, extendibility, minimal encoding bias, minimal ontological commitment [Gruber 1995]. This work and the proposed evaluation framework is largely motivated by the need to contribute to the evaluation of biomedical integration systems so they can gain wide use in distributed computing environments like the semantic web.
2.2 Information Systems Evaluation

The theory of evaluation is rooted in the twin ideas of accountability and social enquiry [Alkin and Christie 2004]. Thus Cronholm and Goldkuhl [2003] divide information system (IS) evaluation approaches into formal rational, interpretive or criteria based, using respectively goal based, goal free and criteria based evaluation strategies. Formal rational evaluations are largely quantitative processes, usually concerned with technical and economic aspects, employing goal based strategies. These strategies focus on intended services and outcomes, to achieve goals which can be phrased in quantitative or qualitative terms.

Interpretive approaches view IS as social systems with embedded information technology. In the same light, Walsham [1993] argues that IS evaluation should consider not only the purpose for conducting the evaluation and associated factors, but also social context and process, and stress the need to consider evaluation as a learning process for all involved. Goal free strategies are appropriate in an interpretive approach, performed with limited evaluator involvement. In contrast criteria based evaluation use selected general qualities for evaluation where scores from multiple criteria are combined into an overall weighted sum [Walsham 1993].

The most appropriate evaluation approach depends largely on its context [Cronholm and Goldkuhl 2003]. Whatever the approach adopted, Alkin and Christie [2004] argue that evaluation models must consider methods, valuing and use. Methods deal with how the evaluation is to be done, and so focus on knowledge construction. Valuing concerns the role of the evaluator in assigning criteria while use focuses on the purpose of the evaluation. These factors can be seen as the process, people and purpose that Ballantine et al. [2000] identify as drivers of IS evaluation.

IS evaluation is considered as a multi-stage process occurring at different points, in different ways, during the product life-cycle. However, IS evaluation literature and practice have been closely linked with the technical project development cycle. For example, Beynon-Davis et al. [2004] propose a model for IS evaluation that distinguishes between the strategic, formative, summative approaches, and post mortem analysis [done if the project has to undergo partial abandonment - which is not relevant to the comparative integration system evaluation being investigated here]. Strategic evaluation is conducted as part of the planning process for an intended IS. Formative evaluation is a continuous, iterative informal process aimed at providing systematic feedback to designers and implementers, influencing the process of development and the final IS [Kumar 1990; Remenyi and Smith 1999; Walsham 1993]. Summative evaluation, usually done at the end of the project is concerned with assessing the worth of a project or program outcome in light of initially specified success criteria [ibid].

BOIS can also be seen as a type of IS, with systemic features that can be variously evaluated against user requirements. What distinguishes them from general IS is arguably the dynamics of the environment in which they are used. The vast amounts
of biomedical data require ontology integration systems that can capture and represent new structure, processes and functional objects that emerge with data in the two domains. In a comparative evaluation of biomedical ontology integration systems as IS, the evaluator is the proposed user. Neither an interpretative nor a formal rational approach seems as appropriate as a criteria-based approach in which the user’s requirements motivate the evaluation criteria. Such an approach can be undertaken in a strategic, formative or summative evaluation.

3. Deriving the Evaluation Framework

This section explores existing approaches to ontology evaluation and systems theory as key underpinning concepts for deriving the framework.

3.1 Ontology Evaluation Approaches in Biomedicine

The pre-modeling, modeling and post release stages of the ontology development life cycle have been used to categorize ontology evaluation [Hartmann et. al. 2004]. Based on type and purpose, Brank et al. [2005] provide a taxonomy of ontology evaluation approaches as follows:

1) Approaches that evaluate an ontology by comparing it to a golden standard ontology or other representation of the problem domain for which an appropriate ontology is needed [Gomez-Perez 1994; Guarino, 1998; Hovy 2001].

2) Task based approaches that use the ontology in an application and evaluate the quality of results [Porzel and Malaka 2004].

3) Data or corpus driven approaches. These evaluate the congruence of an ontology with a given corpus to determine how appropriate it is for the representation of knowledge of the domain represented by the texts [Brewster et al. 2004].

4) Assessment by humans to show how well the ontology meets a set of predefined criteria, standards, requirements as in Ontometric [Lozano-Tello and Gomez-Perez 2004] and the peer-review based approach [Supekar 2005].

3.2. Multicriteria Approaches to Ontology Evaluation

Multiple criteria approaches deal with the problem of selecting a good ontology from a given set based on defining several decision criteria or attributes. For each criterion, the ontology is evaluated and given a numerical score and an overall score computed as a weighted sum of its per-criterion scores [Brank et al. 2005; Lozano-Tello and Gomez-Perez 2004]. While this approach requires a lot of manual involvement by human experts, it allows a combination of criteria at many levels [ibid].

Ontometric [Lozano-Tello and Gomez-Perez 2004] is a multicriteria decision making method for use by knowledge engineers to select an appropriate ontology among various alternatives or to decide the suitability of a particular ontology for a
A Flexible Approach for User Evaluation

3.2. Selection Criteria

A project using the dimensions of: cost, content represented, language, methodology and software development environment used [Hartmann et al. 2004]. The selection criteria is based a four step analytic hierarchy process (AHP) [Saaty 1980] namely:
1) Decide upon the criteria for selection;
2) Rate the relative importance of these criteria using pair-wise comparisons;
3) Rate the potential choice relative to the other on the basis of each selection criterion, achieved by pair wise comparisons of the choices;
4) Combine the ratings derived in steps 2 and 3 to obtain an overall relative rating for each potential choice.

3.3. Contributions from Systems Theory

General systems theory (GST) is here adopted to explain the emergent properties from the complex and dynamic nature of data in biomedical integration systems [Huang et al. 2004]. GST is “elements in standing relationship, the joining and integrating of the web of relationships to create emergent properties of the whole that are not found in any analysis of the parts” [Von Bertalanffy 1962]. GST explains structure and properties of systems in terms of relationships from which new properties of wholes emerge that are not found among those of elements, and the corresponding behavior of the whole cannot be explained in terms of the behavior of the parts. Concepts from the process of self organization (SO) may be used to extend systems theory. In SO, the internal organization of a system increases in complexity without being guided or managed by an outside source and displays emergent properties which do not exist if the lower level is removed [Gershenson 2006]. In SO the environment is unpredictable and the elements interact to achieve dynamically a global function [ibid].

Engineered systems self organize by adaptation, anticipation, robustness or a combination of these features [Gershenson 2006]. Self-organizing systems, rather than being a type of systems, are a perspective for studying, understanding, designing, controlling, and building systems; the crucial factor being the observer, who has to describe the process at an appropriate level and aspects, and to define the purpose of the system; SO can therefore be everywhere, it just needs to be observed [Heylighen and Gershenson 2003]. Organization is seen as structure that has a purpose. The observer has to focus their viewpoint, set the purpose of the system to see the attractor as an organized state at the right level and in the right aspect in order to observe self-organization - a perspective used to design, build, and control artificial systems. A key characteristic of an artificial self-organizing system is that structure and function of the system emerge from interactions between the elements [ibid]. It is this perspective of self organizing systems that is adopted for this study to explain the emergent properties faced by frameworks for evaluating BOIS.

3.4. Field Study

A mixed method research strategy combining quantitative and qualitative deductive survey was used in the study to identify the dimensions and requirements for a
A Flexible Approach for User Evaluation

meta-ontology of biomedical integration systems. Existing literature and document analysis were used to determine scope and identify common structures, functions, inputs, outputs for such a meta-ontology. Similarly, ontological theories that support knowledge integration in biological and clinical integration systems were described. The meta-ontology acts as the frame of reference for the evaluation framework. The dimensions of the reference evaluation framework were tested for potential user agreement using a field study.

The study population consisted of 450 (four hundred fifty) health care workers. Structured interviews and a survey using questionnaires were conducted to collect data and clarify user needs relevant to biomedical integrated systems. In the survey, potential users of such a system were asked for their level of agreement with characteristics of a meta-ontology for such a system, and its evaluation criteria namely: the scope, purpose, requirements, granularity, inputs, processes, outputs.

A high level [>70%] of respondents agreed that a biomedical integration system should:

1) Integrate data from diverse sources and yet useful to molecular biologists, clinicians, healthcare workers, legal practitioners, information systems developers and policy makers

2) Support heterogenous data integration across levels of granularity

3) Support functions and processes that link and relate both existing and future genetic to clinical patient data objects, determine the fit between task/user needs to existing biomedical integration models and provide feedbacks for continuous improvements of such models.

4) Provide for a generic meta-model that define structures (concepts) that are common to both patient clinical and genetic data.

The results of the field study were incorporated into the reference framework. A new framework for evaluating ontology based integrated systems was the result. Inputs, processes, outputs and feedbacks for such a framework were identified from the field study. It relates user needs to structure, functions and processes of an ontology and seeks to determine the effectiveness of a biomedical integration system using completeness to measure system quality.

5.5. The Framework

Existing ontology based evaluation models are attempts to define standards for structuring, representing and integrating existing knowledge in a domain of discourse using specified criteria. Such criteria are often combined and used for user assessment and selection of an ontology appropriate for a given task [Lozano-Tello and Gomez-Perez, 2004; Supekar 2005]. In biomedicine, where new data objects emerge and new user requirements emerge with the rapidly increasing data, flexible models that enable users to iteratively search through an ontology library using multiple criteria are more likely to result into: (1) selection of an
appropriate ontology for a given task or (2) re-specification of new requirements for an ontology to fit the task. This framework is an attempt to evaluate biomedical ontology integration systems as information systems with the evaluator as the proposed user. A criteria-based approach in which the user’s requirements motivate the assessment criteria in a formative evaluation [Kumar 1990; Remenyi and Smith 1999; Walsham 1993] framework is found appropriate as it aims to provide systematic feedback to designers and implementers, influencing the process of development and the final IS.

This framework (in figure 1) extends Ontometric [Lozano-Tello and Gomez-Perez 2004], a multicriteria decision making method for selecting the most appropriate ontology among various alternatives or for deciding the suitability of a particular ontology for a project. Ontometric, based on the AHP framework [Saaty 1980] allows ontology engineers and domain experts the flexibility to select the hierarchy for the decision criteria to be used in evaluations. It however offers no specific features that support evaluating systems in dynamic environments like biomedicine that require reusing, extending and modifying existing ontologies to accommodate new types of data objects and avoid the huge effort of starting or building entirely new ontologies [Alani and Brewster 2006]. The evaluation framework as proposed in this study (figure 1) mitigates this shortcoming by adopting a systems approach so that domain or task evaluator (user) needs are considered when pruning, extending or modifying an existing ontology.

Inputs for the ontology based biomedical evaluation framework are defined in terms of scope, function, structure, clinical and biological objects and their relationships. User requirements for the integration system are the other type of requirements for evaluation. Biomedical ontology integration systems (Ontologies) also have taxonomic structures (defined using objects and relationships between them), function and processes. The processing requirements for the evaluation framework are: determining evaluation criteria from user needs; visualization and identification of relationships between clinical and biological concepts; comparison of the overlap between user needs and the ontology model; determining the fit between a task and user integration requirements; providing feedback to improve upon and enable existing models meet user task requirements. The evaluation process determines the relevance of an integration ontology for a give task. The overlap between user needs and the ontology model determines the fit between the two.
The evaluation framework seeks to establish how well the biomedical integration system conforms to task derived user requirements i.e. completeness. On evaluation, the scope, structure (objects and relations), processes and functions of a given ontology may meet user requirements leading to its selection for use in integration of clinical and biological data objects. Alternatively, a proportion of total user requirements defined may not be met and are here referred to as emergent requirements (ER). Emergent requirements are used as feedbacks for re-examination and reclarification of the scope of user needs, and for reselection or re-specification of the ontology model. This leads to new ontology integration systems with more complex structure and function. General systems theory and that of self organization help to explain these feedbacks caused by these emergent properties of the framework, leading to complexity.
System quality, as used to evaluate an ontology in this framework is defined as “conformance to requirements” [Crosby 1979; Garvin 1984]. A need can be decomposed into a set of requirements upon features and characteristics of a product or service, and if all these requirements are conformed to, then that need has been satisfied. The evaluation criterion seeks to establish how well the biomedical integration system ontology conforms to requirements. This is completeness for the system, and is here expressed as:

\[
\text{Completeness} = \frac{\text{All Requirements met}}{\text{Total system requirements defined}} \quad \text{Equation 1}
\]

In the framework, this relationship expresses how well the ontology based structures satisfy task requirements. Following evaluation, a proportion of total requirements defined are not met. These emergent requirements [ER] require re-examination of user needs and/or ontology structure leading to a redefinition of the integration requirements.

\[
\text{ER} = [\text{Total system requirements defined} - \text{All Requirements met}] \quad \text{Equation 2}
\]

The framework guides users (ontology developers or domain experts) to select an appropriate ontology integration system for a given task through the following steps:

1. Determine the selection criteria (in terms of scope, function, structure, clinical and biological objects and the relationships between them) for an ontology based on the requirements of a given integration task.
2. Rate the relative importance of these selection criteria to the integration task.
3. Select and Visualize ontology to assess (scope, structure, function, processes).
4. Assess ontology by matching it to identified selection criteria for the task.
5. Identify any unmatched (emergent) requirements.
6. Use emergent requirements to help rescope/reselect user or ontology or select matched ontology for use in the integration process.

4. Conclusions

The literature shows the important need for evaluation of ontologies use in integrating data in distributed computing environments. The paper identifies the challenges faced by existing frameworks in evaluating such ontology based biomedical integration systems in a dynamic environment. This framework as given in this paper extends the OntoMetric evaluation method [Lozano-
Tello and Gomez 2004] using systems theory. Systems theory is then adopted as an underpinning theory in a user driven criteria based formative evaluation framework for biomedical ontology integration systems. Steps for the utility of such a framework are also given. The framework emphasizes the properties users tend to look for when judging the suitability of a biomedical ontology integration system. Interviews and questionnaires in a descriptive survey were used to validate the framework. The framework enables users to assess and select a suitable ontology for their particular use case when accessing ever-increasing libraries of biomedical data. The novelty of this framework lies in the ability to relate ontology structure and user derived objectives in order to derive requirements for evaluating ontology based integrated biological and clinical information systems in environments where the amounts of data are ever increasing and user needs are changing.

References


