Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 38 (2019) 1678–1682

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2019), June 24-28, 2019, Limerick, Ireland.

Metrics to gauge the success of a manufacturing ontology

David Koonce, Dusan Sormaz

Ohio University, Athens, Ohio, 45701, USA

Abstract

Ontologies are structures of concepts that define a high-level representation of a system or area. They can serve as a foundational understanding for developing or integrating software representations of said system. And while the construction of can be a complex, multi-party exercise, the assessment of ontologies is often defined less on usage and more on completeness and coverage. In manufacturing, ontology development ranges from supply chain to production to design. Owing to the computer science foundations of ontology design and representation, the value or quality of an ontology can be assessed on notions of completeness and coverage. Recently, researchers have posited that usage should factor into the Ontology Lifecycle. Similar to how the market, and not technology, defines the success of a product or technology, this paper will examine how utilization of an ontology can define the value or quality of the ontology

Keywords: Ontology Evaluation; Ontology Usage; Manufacturing Ontology

1. Introduction

Ontologies are a growing interest in areas that struggle with multiple data formats. Data modeling, the variety of models that exist within a domain and the traditional interoperability of database systems is a complex and time-consuming issue in industry that needs a unified model to support integration [1]. By seeking to lay a foundation of real-world to entities and relationships, an ontology can become a neutral definition that guides the creation, transformation and integration of data within a domain.

Beyond the insight gained from an explicit modelling of a topic of interest, ontologies can support the modelling of data for storage, processing, translation or integration. And, the rich nature of an ontology can support this in

10.1016/j.promfg.2020.01.116

traditional relational or object-based formats. An ontology can be used as a foundation for a semantic analysis of text.

However neutral the intent of an ontology modeller might be, there is an implied perspective that a modeller (or committee of modellers) imparts on their design. Researchers term this a Concept Orientation, and it stems from their understanding of the terms used in the ontology. [2] The bounds of an ontology will also be guided by the purpose of the ontology. An ontology to support manufacturing planning will surely not have types or entities for viruses. And, an immunology ontology will be disjoint to a manufacturing planning ontology. And while an ontology within a domain might be built in as neutral a format as possible, the Concept Orientation and purpose will lead to differences.

But, when multiple ontologies exist in a particular domain which ontology to use may be unclear. Conventional measurers for ontologies exploit the graph properties of ontologies by using common graph measures [3]. However, these may not account for uniqueness in the ontology language, Concept Orientation or purpose. Likewise, if data translation or integration is needed for an ontology-based system, metrics that account for generality or adoption might also be important. Ultimately, the selection of an ontology may be limiting to system development, deployment or integration. This paper summarizes some common approaches to measuring ontology usage and presents an adaptation of these techniques to account for adoptions in the market and the maturity of the ontology. The breadth and scope of an ontology under consideration and its maturity can inform the ultimate decision.

2. Background

2.1. Ontologies

Ontologies define a representation of reality for some domain of interest in a way that different persons can understand that reality [2]. With a combination of types/concepts/entities and their relationships, an ontology can serve as the basis for determining an appropriate data schema, interpreting text and data, and designing and building systems. And, many representational languages and structures exist to develop and present an ontology [4]. With the assumed neutrality of an ontology can aid in the integration of systems and their data. Ontologies are large, complex definitions that can take significant time and resources to construct. Assuring the neutrality of an ontology also requires that multiple parties contribute, in an evolving design process, to the full definition.

Building an ontology can be modelled on the software development lifecycle [4], beginning with managerial activities, then development, and finally support. However, they note that most of the approaches to ontology creation are focused on the development aspects and not as much the lifecycle and management activities.

2.2. Measurement of ontologies

Given the complexity of developing an ontology, researchers have developed techniques and tools for assessing the quality of a given ontology. These techniques centre on the representational structure and completeness of the ontology. Brank and Grobelnik [5] classify ontology evaluation techniques into 5 categories. At their lowest level, (1) Lexical/Vocabulary evaluation focuses on the concepts and data looking at similarity with metrics as detailed as string similarity. By adding (2) Taxonomic and Semantic information, a higher level of evaluation can match similarity of structures. These techniques can employ probabilistic models to assess the quality of a concept match. As ontologies can reference or build off of classes in other ontologies, (3) Context Level evaluations leverage these linkages to score or assess an ontology. Given that an ontology will likely be used in some application, (4) Application-Based evaluations seek to assess the application's success or quality of output. In a more independent analysis, semantic indexing of an ontology and set of domain terms can be used to provide a (5) Data-Based measure of an ontology. These frames for evaluation may stand independently, or be combined into multi-criteria evaluations of an ontology [5].

2.3. Measuring ontology usage

In their book on measuring ontology usage, Chang et. al. present a pathway to measuring and analysing ontology usage that extends beyond conventional measures [6]. In their initial Identification Phase, they present two statistics for ontology usage that denote the identified usage of an ontology. Their Degree Centrality measurement of an ontology sums the connections that ontology shares with other ontologies in the set of ontologies under investigation [6]. This measure aligns with Brank and Grobelnik's Context Level evaluation of an ontology by capturing the linkages between all systems in the domain. [5]

$$C_D(o_i) = d(o_i) = \sum_{j=1}^{n_j} A_{ij} \tag{1}$$

where i = 1,…n_j, or number of nodes/ontologies, n_1 = |O| or ontology, and A_{ij} represents the affiliation matrix for nodes i and j.

And, their Normalized Ontology Usage Degree metric scales Degree Centrality by the data sources under investigation. [6]

$$C'_D(o_i) = \frac{d(o_i)}{number\_of\_datasources - 1} \tag{2}$$

Together, these two measures form a foundation for additional usage analysis and steps. Key to their analysis are questions of interest for ontology owners, data publishers and application developers. These questions focus on who and how many have adopted an ontology, what concepts are used, as well as the level and complexity of their use.

3. Selecting from Competing Systems

Choosing between multiple alternatives is often a non-trivial task. While one alternative may better match needs in one area, another may meet other needs. Or, limitations or costs may be hard to justify if another alternative is "close" to meeting a need. In selecting a software solution, it is often posited that whatever is chosen will be seen as the wrong choice in the future. Key to selecting the best match for systems, like ontologies, is a structured, quantitative process that uses criteria critical to the success of the intended system.

Böhmer et. al. present an approach to seamless interoperability for logistic business objects [7]. In their work, they discuss the challenges of selecting a base object standard for e-business integration. While the selection process is not fully enumerated, the criteria that led to the selection of a business object system were: an extensible standard based on common standards, standard maturity, adoption and use in real system, matching of representation to need, and technology neutrality.

Liu et. al. present a comparison, and via it a set of comparison criteria, for Collaborative Business Process modelling tools [8]. Their comparison framework considers coverage, storage/persistence, deployment, and human tasks.

Where system maturity may conflict with shifting popularity, quantitative models may still prove useful in selection. As an example, the choice of a programming language should involve a set of must-have and want-to-have criteria [9]. These want-to-have criteria then drive the choice amongst tools which meet the must-have standards.

Crucial to any decision for a software or systems platform is the adoption of the technology in the industry or field of interest. Technology adoption models point to the phases in acceptance and adoption of technology and point to how adoption accelerates as a critical mass is achieved. When looking at infrastructure technologies to support collaboration for system design integration, there are several advantages to well adopted technologies. Rocco et.al. [10] point to the opportunities for: a community of users, the availability of models for learning and validation, tools to support interoperability, and collaborative modelling. A common framework leverages interoperability with familiarity to support modelling complex systems that are design-level compatible with other systems.

4. Metrics for Ontology Selection and Success

4.1. Selection of an ontology and success

The widespread adoption of an ontology in a particular domain can then be seen as a generalized success for that ontology, as it works towards a critical mass of adoption. As the utilization count of a representational definition grows in magnitude and proportion, the standardization of that representation grows. Given that an ontology under consideration for adoption meets the “must-have” criteria stipulated in an identification phase, there are 2 additional measures of the suitability that should be considered in an adoption decision.

4.2. Usage metric

The prior adoption of an ontology points to its suitability to another usage. The choice of an ontology by others points to both the identified suitability of the ontology and the relative usage or market share of the ontology. Chang et.al. [6] presented the Normalized Ontology Usage Degree (see Eq. 1) as a measure of the centrality of an ontology by the proportion of other ontologies which link to or share that ontology. Expanding that notion, one metric to the selection of an ontology would be the number of ontology applications which use that underlying ontology. The Usage of an ontology (Eq. 3) is then defined as the sum of all usages observed of the ontology in the set of identified applications. And the Scaled Usage (Eq. 4) is simply the Usage divided by the number of applications.

$$u(o_i)=\sum_{j=1}^{n} o_{ij} \tag{3}$$

where o_i is the ontology, j = 1, … n are the identified applications, and o_{ij} is the usage of the ontology in that application.

$$u'(o_i)=\frac{o_i}{n} \tag{4}$$

The Scaled Usage, or proportion of adoptions, is a unitless measure that shows the likelihood that another, unseen application would utilize the ontology. One may argue that Usage penalizes newer technologies, as fewer applications necessarily exist which have adopted the ontology. To account for that, a measure of the time since inception can be used to account for the maturity of the ontology.

4.3. Maturity/persistence metric

The choice of adopting a specific technology can be cast on the technology adoption cycle of that technology. Early in a technology’s life, adoption carries the risk of low or stalled adoption by others, changes as the technology matures and overall higher adoption costs. Early adopters therefore may expect higher implementation costs and greater risk of failure or abandonment. By looking at the length of use of adoptions, accounting for the usage, a measure of the maturity of a potential technology can be determined. A relatively young ontology with a higher utilization may point to an accelerated technology adoption and a potential reduction in risk. Likewise, an established ontology with a low utilization may point to the ontology following a slower adoption timeline.

The Maturity of an ontology (Eq. 5) under consideration can be defined as Usage of an ontology divided by the number of years an ontology has been available for adoption.

$$m_i=\frac{u(o_i)}{(age)_i} \tag{5}$$

where o_i is the ontology, $u(o_i)$ is the Usage (eq. 3) and $(age)_i$ is the chronological age of the ontology in years.

A larger value points to a greater adoption, relative to age.

5. Conclusions

The growing interest in using ontologies as a foundation for designing systems has led to alternatives for developers. When challenged with a set of choices, with ostensibly comparable modelling richness, the decision maker needs to identify as many decision criteria as possible. Conventional metrics for measuring an ontology rely on graph theory. Newer measures look at usage and the linkages between ontologies to assess the capability and suitability of an alternative. This paper presents two additional metrics to aid a decision maker. The Usage, which captures a market share of an ontology, is a simple proportion of applications using the ontology. And, the Maturity is a measure of the Usage over the life of an ontology. Together, these metrics help guide a choice by establishing a tie into the market of other, historical adopters.

References

[1] R. d. Virgilio, F. Giunchiglia and L. Tanca, Semantic Web Information Management, (2010).
[2] R. Arp, B. Smith and A. D. Spear, Building Ontologies with Basic Formal Ontology, (2015).
[3] D. Vrandečić and Y. Sure, How to Design Better Ontology Metrics, ESWC '07 Proceedings of the 4th European conference on The Semantic Web: Research and Applications, (2007).
[4] O. Corcho, M. Fernández-López and A. Gómez-Pérez, Ontological Engineering: What are Ontologies and How Can We Build Them?, Ontological Engineering: What are Ontologies and How Can We Build Them? (2007), 44-70.
[5] J. Brank and M. Grobelnik, A Survey of Ontology Evaluation Techniques, Proceedings of 8th Int. multi-conf. Information Society 2005.
[6] E. J. Chang, O. K. Hussain, J. Ashraf and F. K. Hussain, Measuring and Analysing the Use of Ontologies: A Semantic Framework for Measuring Ontology Usage, Germany: Springer International Publishing, 2018.
[7] M. Böhmer, M. Schmidt and N. Weißenberg, Seamless Interoperability in Logistics: Narrowing the Business-IT Gap by Logistics Business Objects, Cloud Computing for Logistics, 2015.
[8] H. Liu, Y. Lembaret, D. Clin and J. Bourey, Comparison between Collaborative Business Process tools, 2011 Fifth International Conference on Research Challenges in Information Science, Gosier, FR., 2011.
[9] S. J. Sherman, R. F. Shehane and D. W. Todd, Quantitative Model for Choosing Programming Language for Online Instruction, Journal of Instructional Pedagogies, 20, 2018.
[10] J. D. Rocco, D. D. Ruscio, L. Iovino and A. Pierantonio, Collaborative Repositories in Model-Driven Engineering, IEEE Software, 32, 3, (2015) 28-34.