A Proposed Framework for the Analysis and Evaluation of Business Models

JEAN-PAUL VAN BELLE
University of Cape Town

This paper proposed and validates a new framework for the evaluation and comparison of enterprise models. After a broad literature survey, a large number of evaluation criteria were established and consolidated into a comprehensive framework. The main structuring principle of the framework is the dimension distinguishing syntactic, semantic and pragmatic criteria. The resulting framework is validated against some conceptual principles.

Categories and Subject Descriptors: I.6.4 [**Model Validation and Analysis**]; D.2.8 [**Metrics**]: Product Metrics.
General Terms: theory, design, measurement, verification.
Additional Key Words and Phrases: model analysis, model complexity, CASE measures, quality.

1. BACKGROUND

The development of most large business information systems starts with a domain analysis; this is referred to as *model-driven* development. As many of the programming and implementation activities become automated, the importance of the modelling phase grows. Simultaneously, the move from functional "silo" applications to enterprise-wide systems has increased the scope of the domain models to integrated enterprise-wide models. This move is amplified by the desire to build data warehouses to satisfy the managers' need for integrated information as well as the desire of IS managers to develop long term strategic plans based on, *inter alia*, an enterprise information architecture. Modelling in general has become serious business, of which enterprise modelling is seizing a significant and growing piece.

Academics and practitioners alike have been quick and prolific in developing development tools and methods. Consequently, there have been numerous studies, frameworks, evaluations and comparisons of development methodologies and tools. However necessary these studies of the relative merits of various development tools are, it still is the *output* or *product* which is most important from a business point of view. At this point, there seems to be a particular dearth of guidance available on how to evaluate the actual output of the modelling activity: what measures or techniques exist to evaluate a completed enterprise model? The quality of a model will have a significant impact on the IS function, regardless of whether the model is used as an enterprise data standard, the blueprint for a strategic information architecture, the meta-model for a data warehouse repository or the actual development of an information system.

2. RESEARCH OBJECTIVE

The purpose of this paper is ***the development and theoretical validation of a comprehensive framework for the analysis and evaluation of enterprise models***. No satisfactory framework for enterprise model evaluation was found in the literature, although a number of candidate frameworks for the evaluation of other "intellectual products" exist. These are used as inputs to guide the building of an integrated and comprehensive framework. This paper details some of the theoretical aspects of the framework. An empirical validation of the framework is described elsewhere (Van Belle, 2003).

The framework includes intrinsic qualities (absolute measures that can be computed for one specific model) and comparative qualities (relative measures that compare models). Some of these entail a ranking or judgement (better, worse) whereas other measures merely differentiate (e.g. model A is more like model B, whereas models C, D and E form a separate family). Because of practical and methodological reasons, the emphasis was on static models. However, a parallel and completely independent research effort developed a similar framework to cater for dynamic models (Taylor, 2003). And, although the framework focuses on enterprise models, it should be possible to use the framework for the evaluation of models in other domains, e.g. embedded systems, or sub-domains within the enterprise, e.g. a specific functional area.

Finally, the framework was developed from an interdisciplinary perspective. Since enterprise models themselves originate from such diverse sources as systems theory, computer science, ERP, accounting, linguistics and systems engineering, the proposed framework sought to incorporate contributions from those areas and others, such as construction architecture, complexity theory and aesthetics.

Author Address:
J.P. Van Belle, Information Systems Department, University of Cape Town, P Bag, Rondebosch, 7701, South Africa; jvbelle@commerce.uct.ac.za.

3. PRIOR RESEARCH

A rich body of literature exists on evaluation criteria for models and related conceptual structures. These are found in a variety of reference disciplines: methodology engineering, systems engineering, ontology research, modelling, etc. Criteria can be arranged in a flat (i.e. unstructured sequential) list, in a hierarchical tree structure or in the form of a framework based on some theoretical structuring concept or theory.

Many authors have suggested different lists of criteria to evaluate models (Benyon, 1990; Crockett, 1991, Claxton & McDougall; 2000; Halpin, 2001). Some additional criteria may be gleaned from those that define high-quality data, such as those proposed by Orli (1996) and Courtot (2000). Ontology researchers have also proposed their own criteria for enterprise models (Fox, 1998; Valenta, 1996; Swartout, 1996; van Harmelen, 1999; Noy, 2001). Williams (1996) lists 45 requirements for enterprise reference architectures, whilst Korson & McGregor (1992) propose criteria to evaluate OO models.

A few authors go beyond unstructured lists and organise their evaluation criteria into frameworks, often in the context of evaluating the quality of modelling approaches and methodologies. Structures which organise the different criteria can be matrix presentations, such as the software engineering quality matrices: McCall's quality factors, the very similar Böhm model and Gillies' hierarchical quality model (Böhm, 1976; Gillies, 1997).

Finally, some frameworks are based on an explicit theoretical organising principle. Avison & Fitzgerald (1995) developed a framework based on previous literature but taking contextual and philosophical factors into account. Brazier (1998) has a more structured and detailed framework. The Seligman framework has an interesting process view and formed the basis for Hommes (1998) framework as well as (Van Belle, 2000).

There are a number of problems with the frameworks mentioned earlier. These comments are not intended as criticism because some of the frameworks are taken out of the context for which they were developed.

- ***The Grounding Problem.*** All lack an underlying theoretical basis: although the distinctions may be based on a natural grouping or structuring of the desired factors but there is no underlying theoretical or philosophical basis for the framework dimensions. Although it is acknowledged that the frameworks may still be valuable and valid – as long as they are based on the principles of soundness and completeness – some authors (e.g. Frank, 1999) stress the value and importance of a strong theoretical grounding for an evaluation framework.
- ***The Partial Applicability Problem.*** Most frameworks apply only partially to the evaluation of enterprise models. A lot of criteria relate to the development process, which is not applicable in our case. This is not an intrinsic criticism but refers to the lack of a suitable framework which has been developed specifically in the context of modelling.

4. THE PROPOSED FRAMEWORK FOR EVALUATING THE QUALITY OF ENTERPRISE MODELS

The proposed framework will be developed in the following sequence. Firstly, the two major dimensions, or the conceptual structuring principles, will be explained and motivated. Next, the framework will be "populated" with criteria. Finally, the selected criteria will be cross- referenced against the existing literature.

4.1 Framework Dimensions or Structuring Principles

The first and major dimension is grounded in linguistics, information and communication theory and semiotics. The key distinction used in the framework is the fact that all models – or indeed any informational object – have a syntactic, semantic and pragmatic aspect (see Table 1).

- The *syntax* refers to the type of constructs and the legal ways of combining them i.e. the physical or logical representation of the information. In a model, this refers to the meta-model entities. A syntactic model analysis will therefore be concerned with the number and types of entities, relationships etc.
- The *semantics* refers to the meaning i.e. the sense of the information by interpreting the token or signifier.
- *Pragmatics* refers to the context i.e. considerations, issues and background information which influences or moderates the interpretation of the information.

Table 1: *Main Classification Dimension of Proposed Analysis Framework.*

Classification concept	Related terms and mappings	Reference disciplines
Syntax	Symbols, form, shape, structure	Computer science, software engineering
Semantics	Meaning, denotation, sense	Linguistics, lexicography, information science
Pragmatics	Background, situation, context	Business science, management

The *syntactic* model analysis is concerned with the purely structural aspects of the model, regardless of the underlying meaning of the model and its elements. Most of the analysis will be derived from the hard sciences i.e. the disciplines of software engineering, computer science, graph theory and network analysis. This includes a large variety of standard syntactic metrics relating to size, grouping, layering, inheritance structure, and network visualization, as well as some less standard metrics such as interface aesthetics.

The *semantic* analysis of models is concerned with the intrinsic *meaning* of the model i.e. its relationship and mapping to the underlying domain reality it is representing. The syntactic analysis is content to deal with the structural relations, i.e. shape and form of the entities and their relationships and groupers, and therefore treats all entity and relationship names as mere “alphanumeric labels”. The essence of semantic analysis is to unravel the *meaning* of the name (label, word, token) used for a specific model element (entity, relationship, grouper). Put another way, semantic analysis is concerned with the correspondence (mapping, projection, validity) between the model (abstract or intellectual construct) and the underlying domain (reality). Whereas syntactic analysis is fairly technical and easy to automate, semantic analysis involves the more tricky matters of meaning and interpretation and thus lends itself not quite as easily to objective and/or automated analysis. The reference disciplines for this type of analysis are linguistics and lexicography, while much of the analysis concentrates on similarity, correspondence and cluster analysis.

Pragmatic model analysis, as defined in the framework used for this research, is concerned with those metrics and criteria which cannot be assessed purely on the basis of the information contained within the model, but which requires the consideration of information regarding the use, environment or context of the model i.e. information outside the model. The analysis techniques falling under this heading include the face validity, degree of use, authority of model author, availability, cost, flexibility, adaptability, model currency, maturity and support. Most analysis relies on the searching and ranking of certain specific information details, often involving a degree of subjective interpretation and an understanding of commercial business issues.

This distinction has been around for well over a century and has been well validated in a number of disciplines. Quite a few of the researchers who have been involved with the development of information theory and the philosophical foundations of information systems, have referred to this distinction (Benyon 1990, Stamper 1987). It is therefore very surprising that this has not been applied as such to evaluation frameworks in the field of models, which are “complex information products” *par excellence.* As will be shown in the rest of this research, the classification system is particularly useful since it allows one to classify the metrics for the quality factors as well.

As a second organising principle it is proposed that, within each category above, criteria can be arranged from “absolute” to “relative” measures (Table 2). For some criteria factors, an absolute value can be established i.e. “more is better” whereas for others, no ideal value can be determined *ex ante*, since the desired target value depends on factors not directly related to the intrinsic nature of the model. It must be recognized that this distinction is not always as clear as it may seem: the distinction is actually more a matter of degree and there is a fuzzy cross-over area in the middle where it may be more difficult (or arbitrary) to decide whether a given quality measure is absolute or relative.

Table 2: ***A Possible Second Dimension for Proposed Analysis Framework.***

Classification concept	Related terms and mappings
Absolute measures	Theoretical; “das Model an Sich”; the model as object; objective standards; intrinsic qualities; technical factors; “Conforms to specification”; computer science; academic.
Relative measures	Applied; “das Model für Uns”; the model as subject; subjective standards; extrinsic qualities; business factors; “Fit for purpose”; information systems; practitioner.

Note that, unlike the first dimension which is categorical, the second dimension is continuous.

4.2 Populating the Framework with Detailed Criteria

Table 3 lists the proposed model evaluation criteria within the framework structure. Criteria are grouped into conceptual clusters, which can be sub-divided if one wishes.

It is important to note that the criteria listed here are *simple* evaluation criteria. A criterion that doesn’t seem to fit into any one category, is likely to be a *composite* criterion. For instance, the overall “quality” of a model cannot be fitted into any cell of the proposed framework because model quality represents a composite concept including many of the above criteria, with a weighting very much dependent on the actual purpose of the analysis.

Table 3: ***Populated Framework for Model Analysis.***

	Syntactic	Semantic	Pragmatic
Absolute	Size	Genericity: universality & technical independence	Validity: authority & user acceptance
↑	Correctness; error-free; integrity; consistency	Completeness (domain coverage); conciseness; efficiency	Flexibility; expandibility; portability; adaptability
	Modularity; structure; hierarchy.	Expressiveness Similarity and overlap with other models	
↓	Complexity; density	Perspicuity; comprehensibility; understandability; self-descriptiveness	Price; cost; availability Support
Relative	Architectural style	Documentation	Purpose; goal; relevance; appropriateness

Similarly, there are other "quality-like" evaluative concepts relating to models, which are actually composite concepts consisting of many of the criteria listed in the model and are therefore not used to populate the proposed framework cells. An example is *usability*, which includes all of the pragmatic and most of the semantic and syntactic criteria. Another example is model *dependability*, as discussed Barbacci (1995).

The various criteria are drawn from many sources (see appendix). These sources often ascribe different meanings to certain criteria and, conversely, different authors sometimes use different terms to describe a similar concept. To indicate this overlap in meaning, criteria were grouped into "clusters".

4.3 Mapping the Framework to Earlier Research

It is extremely interesting to note that Fabbrini *et al* (1998) may have come tantalizingly close to suggesting the above framework. They used the same distinction between semantics, syntax and pragmatics in the context of assessing the quality of a software architecture, but interpreted the terms quite differently and did not operationalize their approach in any practical way.

Intriguingly, Brinkkemper (1996:214) makes reference to the doctoral research of John Venable whereby, in his search for a theoretical basis for the CoCoA methodology, he distinguished his criteria for the evaluation of conceptual data models between those relating to semantic concepts, syntactic constructs and more generic requirements such as views and modularization.

These are by no means the only explicit references of the three categories in the context of information systems analysis: the fairly well-known publications from both Benyon (1990) and Stamper (1987) made references to these distinctions, and most computer science research in formal (programming) languages uses the distinction between syntax and semantics extensively. In particular, Stamper proposes his *semiotic framework* to classify information on six different levels: 3 on the human information level: social world, pragmatics, semantics; and 3 on the IT platform: "syntactics", empirics and physical world (1995) and maintains that too much of our research focuses on the "syntactic" elements.

Finally, a parallel but independent research effort concerning the quality of process models, uses the same framework distinctions between syntactic, semantic and pragmatic issues (Taylor 2003).

5. THEORETICAL VALIDATION OF THE PROPOSED FRAMEWORK

This section discusses the conceptual validation of the proposed framework. It is concerned with the foundations and internal structure of the framework from a purely theoretical perspective. Although there is no definitive list of criteria, the literature investigated earlier as a source of model evaluation criteria also provides a large number of theoretical quality principles that are discussed elsewhere [reference omitted]

- **Construct efficiency ("Occam's razor") and simplicity**. The main framework dimension uses only 3 distinct yet highly natural distinctions: syntax, semantics and pragmatics. This creates a sparse matrix with only three (or, if the second dimension is included, six) cells, while providing very useful distinctions. The three distinct columns are representative of the three IS references disciplines. Syntactic criteria originate mostly from the exact sciences i.e. computer science and systems engineering, semantic metrics reflect inputs from the linguistic and informational sciences, and pragmatic measures have a distinct commercial or business focus.
- **Perspicuity**: the framework is intuitively comprehendible. It is relatively straightforward to explain the structure of the framework to a novice and there is little doubt about where new criteria can fit in.
- **Coverage and completeness**. Almost all of the criteria found in the large body of the literature could be placed easily and naturally within the framework. The framework is "logically complete" since most conceivable information objects and intellectual products can be analysed from a syntactic, semantic and pragmatic perspective.
- **Orthogonality**. There is no overlap between the categories: there are no "gaps" between the different cells i.e. new criteria, as long as they are not of a composite nature, fit naturally into one and only one of the cells.
- **Extensibility, customisability, robustness and flexibility**. It is easy to add new criteria to, or remove non-applicable ones from the framework without impacting on the overall coherence and usability of the framework.
- **Genericity, universal applicability, portability and reusability**. The framework can, under certain circumstances, easily be applied to other areas. This is considered one of its greatest strengths above all other frameworks but is demonstrated in more detail elsewhere {reference omitted].
- **Formality, objectivity, absoluteness.** The framework is not formally specified and there is no strictly unambiguous set of definitions that fully express the framework. However, there does not appear to be any other formal framework in this research area.
- **Theoretical foundation:** the framework is grounded firmly in semiotics as well as having a solid philosophical basis. This is also considered a very specific strength of the proposed framework.

Overall, it can be concluded that from a conceptual or theoretical point of view, the framework appears to have a fairly high validity. However, the second dimension (the continuum from absolute to relative criteria) appears to be somewhat less valid.

6. OPERATIONALIZATION AND EMPIRICAL VALIDATION OF THE FRAMEWORK

It is not possible to empirical validate the framework directly. Concrete empirical instruments (measures, proxies, metrics) can be chosen or constructed for each of the criteria in the framework. As an illustrative example, Table 4 lists some well-known software engineering analysis metrics against three of the criteria which fall within the first category i.e. syntactic analysis. Similar metrics were found or developed for all other framework criteria. These were then tested using a test bed of twenty enterprise models in an attempt to *empirically* validate each of the metrics as well as to serve as a "proof of concept" of the usefulness of the overall framework. Due to the lack of space, the results of this analysis cannot be reported here but can be found in (Van Belle, 2003a).

7. CONCLUSION

This research has suggested that the proposed framework is a highly productive and valid approach for evaluating enterprise models. As shown in addendum 1, almost all of the evaluation criteria which have been suggested in the literature can easily be classified within the framework. There remain a number of non-classifiable composite criteria, such as overall "quality" and "usability", but exact and consensus definitions of these tend to be fairly nebulous anyway. Even these composite criteria can benefit from the framework since their constituent attributes can be analyzed in terms of the framework.

Table 4: ***Sample Metrics for Syntactic Criteria.***

Criterion	Suggested (Validated) Metric / measures
Size	Entity count, CASE (concept) count and adjusted CASE count
Structure; hierarchy	number of "Is-A" relationships, use of multiple inheritance, number of unique entities in inheritance graph, inheritance graph size, inheritance connectivity density, inheritance depth / hierarchy nesting level, mean depth of inheritance tree, inheritance width, widest level, number of root classes, average number of children (NOC), standard deviation NOC, number of super-classes, reuse ratio and specialization ratio.
Complexity; density	McGabe's cyclomatic complexity; absolute and relative connectivity; average, harmonic mean, maximum and spread of fan-out; DeMarco's data bang ; plot of Fruchterman-Reingold (for similar-sized models) and others (Van Belle, 2003b)

The most useful dimension within the framework is the separation of syntactic, semantic and pragmatic criteria or factors. Each of these set of criteria has a very distinct tone, reflecting a specific paradigmatic approach. Syntactic analysis has a heavy pure science and engineering flavour, drawing heavily from computer science metrics, systems engineering graph-theoretical and even architectural concepts. Semantic analysis relies mainly on lexicography and computational linguistics, as well as more conceptual information sciences such as meta-analysis or information frameworks. Finally, pragmatic analysis, focused on practical business or commerce issues such as support, pricing, purpose, organizational impact etc. The framework thus brings together the basic constituent reference disciplines of information systems.

A second dimension was proposed for the framework, from absolute towards relative measures, but, being more nebulous and ill-defined of nature, it was not a major focus of this research. It can be regarded as a suggestion for future research or elaboration.

The theoretical validation of the framework emphasized its simplicity, perspicuity, completeness, flexibility, extensibility, universality and its sound theoretical foundation.

Apart from the value of the framework to *classify* existing criteria, the framework should also be seen as an ideal way for the *creative design or generation* of new criteria or measures. Indeed, many metrics suggested in (Van Belle, 2003a) were inspired by thinking about model evaluation along the dimensions identified by the framework. It is therefore suggested that the framework could, and should, be applied to other areas of information systems research.

There is still considerable scope for future research and development of the framework. The second dimension should be refined and tested. The proposed metrics and analysis techniques can be validated further – possibly against other model databases – and added to. More emphasis needs to be placed on structural relationships (such as inheritance) and grouper constructs. Finally, the framework should be empirically validated against other modelling domains.

8. REFERENCES

AVISON, D.E. AND FITZGERALD, G. 1995. *Information Systems Development: Methodologies, Techniques and Tools.* London: McGraw Hill.

BARBACCI, M.R.; KLEIN, M. H.; LONGSTAFF, T. A. *et al.* 1995. *Quality Attributes*. Technical Report 95-TR-021, Software Engineering Institute.

BENYON, D. *Information and Data Modelling*. Oxford, UK: Blackwell.

BRAZIER, F.M.T. AND WIJNGAARDS, N.J.E. 1998. A Purpose Driven Method for the Comparison of Modelling Frameworks. In *Proceedings of the Eleventh Workshop on Knowledge Acquisition, Modeling and Management*, Banff, Canada, 18-23 Apr 1998.

BRINKKEMPER, S.; LYYTINEN, K. AND WELKE, R.J. (*Eds*.) 1996. Method Engineering: Principles of Method Construction and Tool Support. In *Proceedings of the IFIP TC8, WG8.118.2 Working Conference on Method Engineering*, Atlanta, 26-28 Aug 1996. London: Chapman & Hall.

CLAXTON, J.C. AND MCDOUGALL, P.A. 2000. Measuring the Quality of Models. *The Data Administration Newsletter* 2000:14. http://www.tdan.com/i014ht03.htm

COURTOT, T. 2000. What to Look for in Packaged Data Models/Databases. In *Proceedings of the Meta Data Conference*, Arlington, 19-23 Mar 2000.

CROCKETT, H.D.; GUYNES, J. AND SLINKMAN, C.W. 1991. Framework for Development of Conceptual Data Modelling Techniques. *Journal of Information and Software Technology* 33,2: 134-142.

FABBRINI, F.; FUSANI, M. & GNESI, S. 1998. Quality Evaluation based on Architecture Analysis. In *Proceedings of the International Workshop on the Role of Software Architecture in Testing and Analysis (ROSATEA'98)*, Marsala, Italy, 30-Jun to 3-July 1998. Available at http://www.ics.uci.edu/~djr/rosatea/papers/fusani.pdf.

FOX M.S. AND GRUNINGER M. 1998. Enterprise Modelling. *The AI Magazine*, Fall 1998: 109-121.

FRANK, U. 1999. MEMO: Visual Languages for Enterprise Modelling. *Arbeitsberichte des Instituts für Wirtschaftsinformatik* (Universität Koblenz-Landau) 1999:18.

GILLIES, A. 1997. *Software Quality: Theory and Management.* Thomson, London.

HALPIN, T. AND BLOESCH, A. 1999. Data Modeling in UML and ORM: A Comparison. *The Journal of Database Management*, 10,4: 4-13.

HOMMES, B.-J. 1998. *Analysing the Quality of a Business Modelling Technique.* Unpublished Masters Thesis, Delft University of Technology.

KORSON, T. AND MCGREGOR, J.D. 1992. Technical Criteria for the Specification and Evaluation of Object-Oriented Libraries. *Software Engineering Journal* 7,3: 85-04.

MARSHALL, C. 2000. *Enterprise Modelling with UML. Designing Successful Software Through Business Analysis.* Addison-Wesley, Reading (MA).

NOY, N.F. AND MCGUINNESS, D.L. 2001 *Ontology Development 101: A Guide to Creating Your First Ontology.* SMI technical report 2001-0880.

ORLI, R.; BLAKE, L.; SANTOS, F. AND IPPILITO, A. 1996 *Address Data Quality and Geocoding Standards.* Unpublished report. http://www.kismeta.com/Address.html

POWEL, A. AND VICKERS, A. 1996. *A Practical Strategy For The Evaluation Of Software Tools.* In Brinkkemper 1996:165-185.

STAMPER, R. 1997. Semantics. Critical Issues. In *Information Systems Research*, BOLAND, R.J. & HIRSCHHEIM, R.A. (Eds) Chichester: J. Wiley, 43-78.

SWARTOUT, B.; PATIL, R.; KNIGHT, K. AND RUSS, T. 1996. Toward Distributed Use of Large-Scale Ontologies. In *Proceedings of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, Banff, Canada, Nov 1996.

TAYLOR, C. AND SEDERA, W. 2003. Defining the Quality of Business Process Reference Models. In *Proceedings of the 14th Australasian Conference on Information Systems (ACIS)*, Perth, 1-3 Dec 2003.

VALENTE, A. AND BREUKER, J. 1996. Towards Principled Core Ontologies. In *Proceedings of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, Banff, Canada, Nov 1996.

VAN BELLE, J.P. 2003. A Framework for the Analysis and Evaluation of Enterprise Models. *Ph D Thesis* submitted to the Dept of Information Systems, University of Cape Town, February 2003.

VAN BELLE, J.P. 2002. Towards a Syntactic Signature for Domain Models: Proposed Descriptive Metrics for Visualizing the Entity Fan-out Frequency Distribution. In *SAICSIT ACM Conference Proceedings*, Port Elisabeth, Sep 2002, pp.19-29.

VAN BELLE, J.P. & PRICE B. 2000. A Proposed Framework for Evaluating Generic Enterprise Models. *South African Computer Journal*, 26: *69-76*.

VAN HARMELEN, F. AND FENSEL, D. 1999. Practical Knowledge Representation for the Web. In *Proceedings of the IJCAI-99 Workshop on Intelligent Information Integration*, Stockholm, Sweden, 31 Jul 1999.

WILLIAMS, T.J. 1996. The Needs of the Field of Integration. In *Architectures for Enterprise Integration.* BERNUS, P.; NEMES, L. AND WILLIAMS, T.J. (eds.). Chapman & Hall, London, 21-31.