A REVISED MODELLING QUALITY FRAMEWORK

Pieter Joubert, Stefanie Louw, Carina De Villiers and Jan Kroeze
Department of Informatics, University of Pretoria, Pretoria, South Africa

Keywords: Systems modelling, Quality, Quality frameworks.

Abstract: Systems modelling quality plays a critical role in the quality of the final system. Better quality systems are one aspect of addressing system failures which are still common today. This research paper studies quality frameworks for systems modelling techniques, presenting a revised framework. Several authors built their frameworks on the Lindland et al. (1994) conceptual model quality framework. Those frameworks are more abstract and static – they do not clearly illustrate the flow of information through the systems modelling process. The proposed framework makes it much easier to identify which quality aspects have to be in place at which points within the modelling process for it to be successful in its purpose. In addition, it creates awareness on issues such as the kind of skills and background knowledge that people, who are involved in this process, need to have.

1 INTRODUCTION

Most of the previous research on model quality involves the development of frameworks with the purpose of illustrating modelling qualities and related aspects.

These frameworks are mostly abstract representations, which do not clearly illustrate the information flow between the relevant actors involved. In this paper a framework is proposed to identify the quality aspects which influence the creating and understanding of system models, as well as the effect of information flow between the involved actors on the overall quality.

2 BACKGROUND

Systems modelling is used to communicate with a number of different people involved in the system development lifecycle. This would include communicating with non-technical users as well as technical support staff, such as database administrators, programmers, testers, etc. of these models. Models are used for communicating the most important statements from the domain, as well as to facilitate the understanding and agreement of problem statements from the domain. They are also used for the development of the system and, should these models not correctly represent the domain, the system will not be adequate and will most likely fail in its purpose.

3 MODELLING QUALITY FRAMEWORKS

Several conceptual model quality frameworks have been proposed of which the Lindland et al. (1994) framework has been used as a foundation by most. The framework identifies three qualities, namely *syntactic quality*, *semantic quality* and *pragmatic quality*. The framework consists of four cornerstones, namely the *domain*, referring to the relevant and correct statements to solve the problem; the *model,* referring to both an implicit and explicit model of the statements actually made; the *language*, referring to the language used according to the specific modelling language syntax; and *audience interpretation*, referring to both technical or social actors who are interpreting the model.

This framework has formed the basis for many extensions and improvements by a number of authors (Siau & Tan, 2005; Krogstie, 1998; Cheng et al., 2001; Krogstie et al., 2006; Jørgensen, 2004; Wand & Wang, 1996; Nelson & Monarchi, 2007; Gemino & Wand, 2004 describing Norman's Theory of Action). All of these frameworks were incorporated in the framework proposed in this paper.

Joubert P., Louw S., De Villiers C. and Kroeze J. (2009).
A REVISED MODELLING QUALITY FRAMEWORK.
In *Proceedings of the 11th International Conference on Enterprise Information Systems - Information Systems Analysis and Specification*, pages 162-167
DOI: 10.5220/0001991001620167

The following concepts are borrowed from the Lindland et al. (1994) framework:

- *Syntactic quality* which refers to the extent to which the model corresponds to the modelling language;
- *Semantic quality* which refers to the extent to which the model corresponds to the domain; and
- *Pragmatic quality* which refers to the extent to which the model corresponds with the audience's interpretation of the model.

The following concepts are borrowed from the Krogstie (1998) framework:

- *Perceived semantic quality* was identified by Krogstie et al. (1998), but has been revised by Krogstie et al. (2006) to say that it should only refer to semantic quality and not perceived semantic quality. The proposed framework in this research paper will, therefore, refer only to semantic quality and not perceived semantic quality;
- *Physical quality* refers to the fit between the modelling language and the participants. It is the knowledge of the participant that is externalised by the use of the modelling language. It refers to both externalisation, which is the level of externalisation of the participant's knowledge, and internalisation, which refers to the knowledge obtained through the interpretation of the model;
- *Social quality* refers to the level of agreement on the model viewer's interpretation of the model. In the newly proposed framework, social quality refers to the level of agreement on all participants' interpretation of both the implicit and external model; and
- *Knowledge quality* refers to the correlation between the participants' knowledge and the domain.

The following concepts are borrowed from the Krogstie and Sølvberg framework, adapted from Siau & Tan (2005):

- *Empirical quality* refers to the 'error frequencies' that occur when the model is created or viewed and consist of comprehensibility matters such as graph layout and readability indexes for text.

The following concepts are borrowed from the SEQUAL framework, adapted from Krogstie et al. (2006):

- *Organisational quality* refers to the correspondence between the model and the organisational goals or earlier base-lined models; and
- *Tool quality* refers to the technical actor's interpretation of the model which occurs through the use of a software tool.

The following concepts are borrowed from the revised SEQUAL framework, adapted from Krogstie et al. (2006):

- *Ideal semantic quality* (prescriptive) refers to the model correspondence between the model and the organisational goals or earlier base-lined models. Ideal semantic quality corresponds with the organisational quality mentioned earlier; and
- *Ideal semantic quality* (descriptive) has the same meaning as the semantic quality, which was used in the previous SEQUAL framework.

The following concepts are borrowed from the Nelson & Monarchi (2007) framework:

- *Inferential quality* will test the reasonableness of the inferences taken from understanding the representation.

4 THE REVISED FRAMEWORK

The revised framework (see figure 1) takes into account all of the previous modelling quality frameworks. Aspects from the framework are as follows:

- *Domain* refers to all possible correct and relevant statements necessary to solve the problem. The domain can also be seen as being synonymous to the ideal knowledge about the domain;
- *Domain expert* refers to the expert who holds knowledge about the domain in a cognitive manner and usually works in the domain area to be modelled. The domain expert may also provide already explicit documentation about the domain to the model creator;
- *Perceived domain* refers to the understanding about the domain which is held by the domain expert cognitively. The knowledge about the domain may be partial or incorrect and therefore refers to cognitive understanding;

- *Language 1* refers to the language which is used to communicate the perceived model of which the outcome is represented as an implicit model. The language is usually natural language such as English.;
- *Implicit model* refers to the model which is created by the domain expert after relevant problem statements have been elicited and communicated from his or her perceived domain view;
- *Model creator* refers to the person who will create the explicit model and is also in some cases referred to as the analyst or the technical actor;
- *Perceived implicit model* refers to the understanding of the model creator about the implicit model which was communicated through Language 1. The model creator should also be able to understand the language which was used to communicate the implicit model;
- *Language 2* refers to the language which is used to create the explicit model. An example of a language is UML for the creation of use case diagrams;
- *Explicit model* refers to the actual explicit model which was created by the model creator;
- *Model viewer* refers to the person(s) viewing the explicit model and includes the model creators, domain experts, as well as other participants. A time aspect is of relevance, because the domain expert is part of the model viewer group, after the explicit model is created; and
- *Perceived explicit model* refers to the understanding held by the model viewer when the explicit model is read. In this case the model viewer also has to understand Language 2 to be able to understand the model.
- *Optimal domain* is the situation that the organisation would want; even though they feel that it is too simplistic to expect all members of the organisation to have the same view of the optimal domain.

It is important to note that the information flow through the modelling process is not strictly sequential, but also involves iterative processes.

Within this proposed framework several modelling stakeholders are represented, namely: *domain experts*, *model creators* and *model viewers*. These stakeholders are subjected to several qualities discussed below (see figure 1).

The **domain experts** are subjected to quality areas, which impact on how well the implicit model is communicated to the model creator. These qualities include the following:

- *Social quality 1* refers to the level of agreement between the different domain experts on the implicit model;
- *Empirical quality 1* refers to the 'error frequencies' that occur in the process to create the implicit model;
- *Syntactic quality 1* refers to the extent that the model is corresponding with the modelling language. Usually the domain experts communicate with natural language to the model creator. If the domain expert's language ability is poor, it will reflect on the syntactic quality;
- *Semantic quality 1* refers to how the implicit model corresponds to the domain. Here, the implicit model is measured against the domain;
- *Knowledge quality 1* refers to the participant's knowledge of the domain. Here it is looked at how well the domain actor's knowledge corresponds with the domain; and
- *Physical quality* refers to the knowledge of the participants that is externalised, indicating the fit between the participants and the modelling language.

It can be seen that several qualities have to be in place when communicating to the model creator. The domain expert also has to understand the model that is created, which will be discussed further under the model viewer section. Both the model creators and the domain experts form part of the model viewer group, which will be discussed in more detail as well.

The **model creators** are subjected to two areas namely the understanding of the implicit model which is created by the domain expert, as well as the creation of the explicit model through a modelling language. The qualities relevant to the understanding of the implicit model are identified as follows:

- *Syntactic quality 2*: The domain expert communicates through natural language and the model creator should be able to understand the language and its syntax;
- *Empirical quality 2* refers to the 'error frequencies' which occur when the implicit model is understood;
- *Social quality 2*: In some instances more than one person may be involved with the creation of models, which means that there could be a

communication aspect between the involved model creators;

- *Knowledge quality 2*: The model creator's knowledge should correspond with the domain while the model is created, otherwise the model will not reflect the problem statements from the domain correctly;
- *Pragmatic quality 1* refers to how well the implicit model created by the domain expert is interpreted by the model creator; and
- *Organisational quality* refers to the correspondence between the model and the organisational goals or earlier base-lined models.

The qualities which are of relevance when the model creator creates the explicit model are identified as follows:

- *Empirical quality 3* refers to the 'error frequencies' which occur when the explicit model is created;
- *Tool quality* refers to the correspondence between the explicit model, created with the use of a software application and the model creator's interpretation of the model;
- *Syntactic quality 3* refers to the correspondence between the model and the modelling language. The modelling language consists of certain modelling notation which is used to create the model;
- *Knowledge quality 2* refers to the model creator's knowledge about the domain when the explicit model is created. The problem statements from the domain should be correctly understood to be able to reflect it correctly on the model;
- *Social quality 2*: When a model is created, agreement between the model creators and the implicit model should first take place before the explicit model can be created; and
- *Semantic quality 2* refers to how well the explicit model corresponds with the problem statements from the domain.

The **model viewers**, also referred to as audience in the literature, include both the domain experts (also referred to as social actors) and the model creators (also referred to as technical actors).

The following qualities are relevant to the understanding of the explicit model:

- *Syntactic quality 4* refers to the extent of which the model corresponds with the modelling language. The model viewers will also have to be trained on the modelling language to understand the notation;
- *Knowledge quality 3* refers to how well the model viewers are knowledgeable about the domain. Their background knowledge about the domain will impact on how the explicit model is perceived;
- *Social quality 3* refers to the agreement between the model viewers and the explicit model;
- *Empirical quality 4* refers to the 'error frequencies' which occur when the explicit model is viewed;
- *Pragmatic quality 2* refers to how well the explicit model corresponds with the model viewer's interpretation of the model; and
- *Inferential quality* tests the reasonableness of the inferences taken from understanding the explicit model.

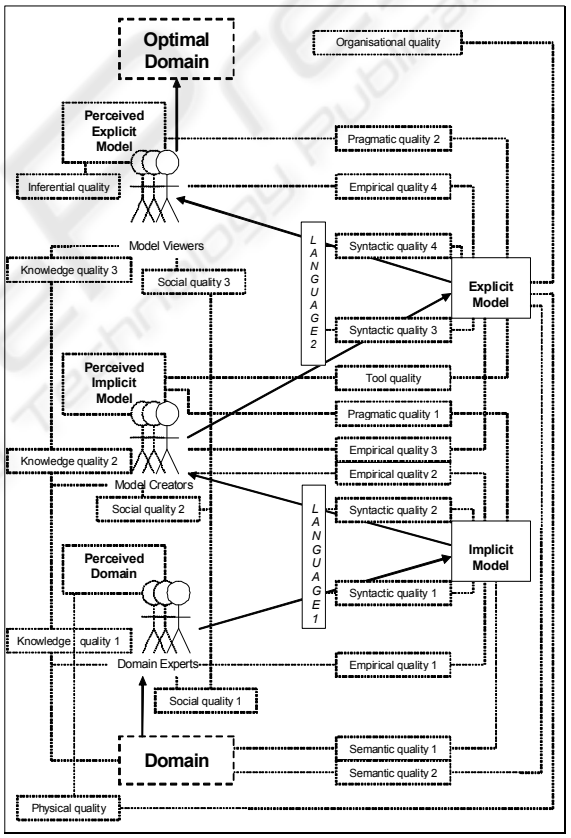

Figure 1: The proposed modelling quality framework.

5 CONCLUSIONS

By looking at the newly proposed modelling framework, a person is able to see the many problematic issues which could influence the modelling process. Quality aspects are illustrated

which all need to be in place for the modelling process to be a quality one. Should all quality aspects be in place except one, the modelling process will be flawed. It has been previously mentioned by Moody et al. (2003) that some of the quality aspects may be more important than others. They concluded that semantic quality is the most important, but it should be noted that they have focused only on semantic, syntactic and pragmatic quality in their study.

Krogstie et al. (2006) have decided not to include physical, empirical and syntactic quality into their revised framework as they feel it is not the most problematic. By contrast, Moody et al. (2005) have identified syntactic quality as an important basis for other qualities, because they say by improving syntactic quality and semantic quality, pragmatic quality will improve, because if the model is not of good syntactic quality it will be difficult to interpret it. If pragmatic quality is improved, semantic quality will be improved, because a model that is difficult to interpret will not be related to the domain.

Several quality aspects are relevant to the actors in the modelling process: the *domain expert*, *model creator* and *model viewer*. The new framework illustrates the information flow through the modelling process, between the different actors.

By using this framework, arguments from the literature can be evaluated accordingly. As an example, Gorla & Lam (2004) indicate that an analyst's (in this section referred to as model creator) analytical skills are more important than behavioural skills in small teams. They say that in smaller teams the analyst may be given additional tasks to systems analysis, which can include system design and programming. In larger teams, the systems analyst may be tasked only with requirement determination and system specification. In the newly proposed framework, more qualities relate to a person's analytical skills than behavioural skills, which would substantiate the argument presented by Gorla & Lam (2004).

Several quality aspects are relevant to the actors in the modelling process: the *domain expert*, *model creator* and *model viewer*. The new framework illustrates the information flow through the modelling process, between the different actors:

Domain expert(s):

- Need to be able to communicate well to each other in order to improve *social quality 1*. Certain organisational communications channels also need to be in place e.g. e-mail, telephone, bulletin boards, etc., to facilitate communication within the group.
- Have to receive training and coaching to improve their knowledge about the domain, consequently improving *knowledge quality 1* as well as *semantic quality 1*.
- Need to be able to communicate through a natural language such as English. If the person's language ability is not good, he or she needs to attend training to improve it, consequently improving *syntactic quality 1*.

Model creator(s):

- Should be able to communicate effectively with the domain expert through natural language, consequently improving *syntactic quality 2*.
- Should ensure that efficient communication channels exist to facilitate communication between them. They need to agree on the implicit model as created by the domain expert. They could discuss and document their viewpoints of the implicit model, which can also be seen as the perceived implicit model, in order to improve *social quality 2*.
- Need to have background knowledge of the domain as well as having the skills to create the explicit model in order to improve *knowledge quality 2*. The skills needed to create the model include knowledge about the language as well as the software tool which is used to create the model.
- Need to possess good analytical skills in order to interpret the implicit model, where after an explicit model is created. If the interpretation of the implicit model can be improved, *pragmatic quality 1* will also be improved. An iterative process of understanding will improve the agreement on the implicit model.
- Should be enabled, by using quality software applications, also referred to as computer-aided modelling tools, to create a model with all needed domain statements, hence improving *tool quality*. Tool quality also has an impact on *syntactic quality*, because with the use of a software application, syntax can be checked. By using a software application, manual checking is less and it is faster to create and check the model. Computer-aided modelling tools may help to limit errors occurring, through automatic layout and model organisation, hence improving *empirical quality*.

- Need to be aware of all organisation goals or earlier base-line models in order to create the explicit model according to it. Background knowledge on the organisation would be needed in order to improve *organisational quality*. Should the model not fulfil the organisation goals, the system will most likely fail its purpose.
- *Need to watch* the correspondence between the domain and explicit model, which would mean that all the qualities in the process up to this point need to be successful in order for *semantic quality 2* to be good.
- Need to have the skill in order to be able to create a model according to the modelling language. If the model is not created with the correct modelling syntax, *syntactic quality 3* will be poor.

Model viewer(s):

- Need to be able to understand the explicit model, therefore knowledge about the modelling language is needed. The modelling process is also a learning process and the model viewers may need training in order to understand the modelling language syntax. This will improve *syntactic quality 4*.
- Need to have background knowledge on the domain in order to improve *knowledge quality 3*. The model viewer may communicate with several domain experts, as well as request any related documentation regarding the domain to get a better picture of the domain.
- Need to communicate with other model viewers in order to agree on the explicit model, hence improving *social quality 3*. Meetings could be held in order to improve communication between model viewers.
- Should be asked to provide their prompt feedback on the explicit model to be able to reach an agreement, improving *pragmatic quality 2*. A process of iteration will also assist to achieve better agreement between the stakeholders. *Social quality* also plays a role here, because if communication is improved between the model viewers, it will be easier to reach an agreement. This can also be seen as an iterative process of understanding, hence improving *inferential quality* as well.

REFERENCES

Cheng, P.C.H., Lowe, R.K. and Scaife, M. 2001. Cognitive science approaches to understanding diagrammatic representations. *Artificial Intelligence Review*, 15(1-2), pp79-94.

Gemino, A. and Wand, Y. 2004. A framework for empirical evaluation of conceptual modeling techniques. *Springer* (9) pp248-260.

Gorla, N. and Lam, Y.W. 2004. Who should work with whom? Building effective software project teams. *Communications of the ACM* 47 (6) pp79-82.

Jørgensen, H.D. 2004 Interactive process models. PhD Thesis, Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim., Norway. *Springer Verlag*: Berlin, Germany.

Krogstie, J. 1998. Integrating the understanding of quality in requirements specification and conceptual modeling. *Communications of the ACM* 23 (1) pp86-91.

Krogstie, J., Sindre, G. and Jørgensen, H. 2006. Process models representing knowledge for action: a revised quality framework. *European Journal of Information Systems* (15) pp91-102.

Lindland, O.D., Sindre, D. and Sølvberg, A. 1994. Understanding quality in conceptual modeling. *IEEE Xplore* (11) pp 42-49.

Moody, D.L., Sindre, G., Brasethvik, T. and Sølvberg, A. 2003. Evaluating the quality of information models: empirical testing of a conceptual model quality framework. *Presented at the 25th international conference on software engineering*, Portland, Oregon.

Nelson, H.J. and Monarchi, D.E. 2007. Ensuring the quality of conceptual representations. *Software Quality Journal* (15) pp213-233.

Siau, K. and Tan, X. 2005. Improving the quality of conceptual modeling using cognitive mapping techniques. *Data & Knowledge Engineering* (55) pp343-365.

Wand, Y. and Wang, Y. 1996. Anchoring data quality dimensions in ontological foundations. *Communications of the ACM* (39) 11 pp86-95.