Towards a Framework for Schema Quality in the Schema Integration Process

Peter Bellström[1] and Christian Kop[2]

[1] Department of Information Systems, Karlstad University, 651 88 Karlstad Sweden
Peter.Bellstrom@kau.se
[2] Institute for Applied Informatics, Alpen-Adria Universität Klagenfurt, Klagenfurt, Austria
chris@ifit.uni-klu.ac.at

Abstract. In this paper we analyze and discuss schema quality in the schema integration process. In doing so, we apply a framework for evaluating the quality of a conceptual schema (e.g. conceptual database schema). In our analysis we combine quality factors and quality metrics with the schema integration process, which is often described as having four distinct phases: pre-integration, comparison of the schemata, conforming the schemata and merging and restructuring. As its main contribution, the paper offers not only new insights on how to improve the quality of the integration process but also a suggestion that the definition of a high quality schema differs between the phases in the schema integration process.

Keywords: Schema Quality, Model Quality, Schema Design, Schema Integration, Conceptual Modeling, Database Design.

1 Introduction

Quality of schemata is very important. We will therefore discuss where quality factors and metrics can be applied in the schema integration process. However, we will mainly focus on integrating structural aspects (e.g. concepts of an enterprise and their relationships to each other). We have therefore adopted quality factors, quality metrics and integration process models from the early conceptual modeling step of database design [1][18] due to the level of abstraction. Particularly, we will describe which of the metrics for quality factors introduced for schema development in general also play an important role during the integration process.

When doing conceptual modeling of schemata for databases as well as for enterprise models, it is important that the stakeholders, e.g. business users and data analysts, first design the schemata for each group of stakeholders (schema design) and then integrate these schemata into one global schema (schema integration). This is vital because the stakeholder schemata not only illuminate differences among user views but also because a global conceptual schema might instead mask these [25]. Schema integration is a complex, time-consuming and error-prone task [24] and is described in [2] as "the activity of integrating the schemas of existing or proposed databases into a global, unified schema" (p. 323). Schema integration not only refers to the process (in [2] the authors view it as composed of four phases), but also the product, as expresed by [9]: “The term integration represents both, a process and its results” (p. 112). The integrated schema (i.e. the product or result according to [9]) should be evaluated according to several quality criterion or quality factors such as completeness, minimality/simplicity and understandability [1][2][20]. However, these quality criteria/factors are not necessary valid for all schemata in all phases in the schema integration process; instead, one criterion/factor could influence another criterion/factor in a negative way [21] or even cause semantic loss [4]. In this paper, we therefore address schema quality within each of the four phases in the integration process. To do so, we apply the framework for evaluating and improving the quality of conceptual schemata described in [20][21][22].

Our research approach can be described as design science [13][29] and our main contribution as a method. By that we mainly mean new insights on how to improve the integration process (method).

This paper is structured as follows: in section two we address related work and in section three our research approach. In section four we address the schema integration process and in section five the applied framework for schema quality and the main contribution of this paper: an analysis on schema quality in the schema integration process. Finally, the paper closes with a summary and conclusions.

2 Related Work

Though quality is a feature of a software product or software artifact, it can be distinguished between quality of the product (artifact) and quality of the process. The quality of the latter of course supports the quality of the product and hence is only introduced for this reason. For the quality of conceptual schemata, a lot of work has been written that examines the quality of the product.

In [1], the authors name a list of characteristics a schema must provide (i.e. correctness, completeness, readability, comprehensibility, consistency, minimality, expressiveness, self-explanation and normality). In [17], the authors subsume this and other research work to a framework consisting of the three dimensions: “syntax”, “semantic” and “pragmatics”. In the [17], the listed characteristics are then related to these dimensions. The syntax-dimension reflects the aspect that a schema must be legal with respect to its vocabulary and grammar (i.e. meta-model). The semantic dimension relates the used terms and notions to the domain context. The chosen notions modeled by modeling elements must be legal and relevant in the domain, and they must be relevant and legal for the purpose for which the model has been built. Finally, the pragmatic dimension introduces the audience, namely the involved stakeholders who have to read and review the schema. A pragmatic quality is achieved if the audience can understand and follow the schema.

After evaluating several quality research papers for a conceptual model, [23] concluded that there is still a need for consensus. What does quality mean? Standards are needed, which are also accepted by the industry. Though standards are necessary, they must nevertheless be adapted to certain issues of conceptual models, ie. which type of model is used (data models, behavior models), and which language is required for a certain type of model (UML diagrams, ER diagrams for data models).

In [22], the authors conducted an empirical study about improving the quality of data models. Particularly, they focused on process quality for the development of data models, which was evaluated in a large Australian bank. Starting with an initial quality model framework that consisted of the model quality factors completeness, simplicity, flexibility, integration, understandability, and implementability, they concluded that integrity and correctness must be added as important factors influencing the quality data model. In the empirical study, it was also important, that the quality was checked throughout the model development process. In particular, quality-checking was not only made at the end of a phase but before, during and after model development phases (e.g. requirements definition, logical design).

In [7], the authors present a metamodel for measured and perceived quality characteristics of a conceptual ER schema. Afterwards, it is evaluated how good measures of four quality characteristics (clarity, simplicity, expressiveness, and minimality) work in practice.

Another framework is the "Guidelines of Modeling (GoM)" [3]. Six principles of modeling are introduced in this framework, namely correctness, relevance, economic efficiency, clarity, comparability and systematic design. These principles can be seen as general strategic and objective definitions for modeling. Based on these goals, the concluded modeling process consisted of the following steps: goal definition (i.e. what is the purpose of modeling), construction of an overall navigation and structural framework (i.e. this navigation and structural framework shall prevent loss in the many models that are constructed), modeling as such, and completion and consolidation.

In the remainder of the paper, we will restrict ourselves to the framework given in [22]. Particularly, we will describe how quality factors and their metrics can be applied in the schema integration process.

3 Research Approach

The research approach adopted within this work can be characterized as design science, see [12][13][14][19]. In design science research, the result is always an artifact, or more precisely stated as "The result of design-science research in IS is, by definition, a purposeful IT artifact created to address an important organizational problem" (p. 82) [13]. Furthermore, the artifact can either be classified as a construct, a model, a method or an instantiation [13][19].

In [13], the authors proposed seven guidelines that researchers should follow to reach good design science research results.

As mentioned in the introduction, the main contribution of this paper is to offer new insights on how to improve the schema integration process, classifying our contribution as a method. This means that we fulfill the first guideline: ***Design as an Artifact***. One problem within the schema integration process is to produce a high quality schema as its end product. In our analysis on schema quality in the integration process, we contribute to new insights on what a technology-based solution, a semi-automatic schema integration application, needs to take into account while integrating several schemata into one global schema. This means that we have fulfilled the second guideline: ***Problem Relevance***. In our analysis, we evaluate our research results using "Informed argument: Use information from the knowledge base (e.g., relevant research) to build a convincing argument for the artifact's utility" (p. 86) [13]. However, we do not claim that we fully have conduct all three cycles in what [15] describe as the relevance cycle, the design cycle and the rigor cycle, since no field testing has yet been carried out. Nevertheless, informed argument has still been used for evaluation and we therefore have fulfilled the third guideline: ***Design Evaluation***. In our analysis, we also give examples and relate how the quality factors could be mapped and used in the integration process, meaning that we have fulfilled the fourth guideline: ***Research Contributions***. We have also combined results achieved in the field of quality of schema development (product quality) with research about the schema integration process in order to give a framework in which integration process step, where a quality measure can be applied. This means that we have fulfilled the fifth guideline: ***Research Rigor***. In our analysis, we have worked iteratively, searching for the best combination of quality factors and quality metrics within the schema integration process. This means that we have fulfilled the sixth guideline: ***Design as a Search Process***. Presenting our results in this paper, we have also fulfilled the seventh and last guideline: ***Communication of Research***.

4 The Schema Integration Process

Our point of departure in this section is the work reported in [2]. In the paper, the authors divide the schema integration process into four distinct phases: pre-integration, comparison of the schemata, conforming the schemata and merging and restructuring. To grasp what schema integration is all about and to have a reference point while discussing schema quality in the schema integration process, each of the included phases with their in- and output will hereafter shortly be described.

4.1 Pre-Integration

Pre-integration is the first phase in the schema integration process. Input to this phase is a set of schemata, so-called user views, that have been designed by the stakeholders. This phase is also one of the least researched phases [26]. In [2] it is pointed out that in the earlier integration methods this phase was often overlooked.

In [26], the author mentions that pre-integration has three main tasks that should be carried out: translating all schemata to the chosen modeling language (canonization), checking for differences and similarities in each schema (intra-schema) and selecting the integration strategy. In [6], the authors proposed three additional tasks that should be carried out in the pre-integration phase: schema element name adoption, schema element disambiguation and introduction of missing relationships.

The output from this phase is a set of revised schemata, the definitions of schema elements and the chosen integration strategy.

4.2 Comparison of the Schemata

Comparison of the schemata is the second phase in the schema integration process. The input to this phase is the output from the following pre-integration phase. Comparison of the schemata has received a lot of attention within the research community and has been mentioned as an important [26] and difficult [8][16] phase. In [15], the author also points out that this phase has three main tasks that should be carried out: recognition of name conflicts, recognition of structural conflicts and recognition of inter-schema properties.

In [2], the authors state that a name conflict can either be classified as a homonym or a synonym conflict. In [1] and [26], one additional conflict was added to the list of name conflicts: reverse subset relationship or cyclic generalization. This type of conflict occurs when e.g. concept A is defined in schema 1 as a specialization of concept B and in schema 2 as a generalization of B.

In [2], the authors also state that a structural conflict can either be classified as a type, a dependency, a key or a behavioral conflict.

The last task to perform in this phase is recognition of inter-schema properties. Inter-schema properties are not really conflicts but instead describe specific dependencies between concepts such as hypernym-hyponym and holonym-meronym dependencies.

The output from this phase is schema element similarities, differences and inter-schema properties. However, this is not the only output since the input to this phase is also forwarded to the next phase due to information that might facilitate the work conducted in the following two phases.

4.3 Conforming the Schemata

Conforming the schemata is the third phase in the schema integration process. The input to this phase is the output from the previous comparison of the schemata phase. This phase has also received some attention within the research community and has been mentioned as the most critical phase [16] and the key issue [27] in schema integration.

In conforming the schemata, the input schemata are adjusted to resolve the recognized similarities and differences. How these similarities and differences are resolved strongly depends on the applied modeling language and whether the schemata are designed on the implementation dependent level or not. For instance, working with schemata on an implementation-neutral level, it is important that the resolution techniques do not delete any modeling elements without being 100% certain that the element is redundant or that it is possible to deduce an element from one or several other elements [5]. The recognized inter-schema properties are studied in this phase. However, the full value is not shown in this phase but instead in the following phase.

The output of this phase is a set of revised schemata, inter-schema properties as well as the input to this phase.

4.4 Merging and Restructuring

Merging and restructuring is the fourth and last phase in the schema integration process. The input to this phase is also the output from conforming the schemata.

In this last phase, the schemata are first merged into one global intermediate schema. The intermediate schema is then restructured since new dependencies might be needed and during this task the recognized inter-schema properties are used as guidance. Additionally, truly redundant schema elements might be recognized and deleted. However, if the stakeholders are not 100% certain that the element is redundant, it should be kept in the schema, as it could result in semantic loss [4]. This task results in a new intermediate schema.

The last task to perform in this phase is to check and verify that the schema fulfills the stated quality criteria [1][2] and/or quality factors meaning, applying a framework for evaluating the schema quality (e.g. [22]). According to our point of view, the research within schema quality that has been conducted until now has focused on this part of the design process and has assumed that the integration process has already been conducted with a satisfying result.

Finally, the output from this phase should be a high quality schema that can be viewed as input for a next, more implementation-oriented phase.

5 Schema Quality in the Schema Integration Process

In this section, we first give an overview of the adopted quality framework developed and described in [20][21][22]. We then analyze how the framework could be applied and adopted in the schema integration process as described by [2]. In doing so, we list each quality factor together with the metrics for each factor as described in [20]. We then analyze, motivate and describe why a specific metric should/could be applied or not for a specific quality factor in a specific phase in the integration process. We also comment on each quality factor as such and how it fits into in the integration process.

5.1 A Schema Quality Framework

The framework for schema evaluation was first proposed in [21] and later revised in [20]. The framework was developed for the ER modeling language and a comprehensive description of the development and evaluation of the framework is described in [22]. The final version of the framework as described in [22] is comprised of six entities: Quality Factor, Stakeholder, Quality Metric, Improvement Strategy, Quality Issue and Quality Review. However, in this paper we will mainly focus on the first three. The authors of [22] also state that Business User, Data Analyst, Data Administrator and Application Developer are the main stakeholders of the schema currently being evaluated. They also conclude that quality factors that improve schema quality are: Completeness, Integrity, Flexibility, and Understandability for Business Users; Correctness, and Simplicity for Data Analysts; Integration for Data Administrators; and: Implementability for Application Developers.

Each stakeholder should be involved in both the schema design and schema integration process. However, it should be noted that Business Users and Data Analysts focus on the earlier parts in schema design and schema integration, while data administrators and application developers focus on the later parts. In this paper we mainly focus on the earlier phase of conceptual modeling meaning, the implementation independent level with schemata that are implementation neutral. It is our opinion that the chosen framework is a good point of departure when researching schema quality of implementation neutral schemata within the schema integration process.

5.2 Schema Quality in the Schema Integration Process

The first and most important quality factor is ***completeness*** [20]. A schema is complete if it contains all user requirements (and nothing but the user requirements). Completeness is measured on the basis of four metrics: number of schema elements that are not part of the user requirements (M1), number of user requirements that are missing in the schema (M2), number of inaccurately defined modeling elements (M3) and number of mismatches with the dynamic/behavioral schema (M4) (see Table 1).

Table 1. Quality Factor 1 *Completeness* and the Integration Process

Metric	Pre-Integration	Comparison of the Schemata	Conforming the Schemata	Merging and Restructuring
M1	YES	NO	NO	YES
M2	YES	YES	YES	YES
M3	YES	YES	YES	YES
M4	YES/NO	NO	NO	YES

Before moving on to each quality factor a comment on the last phase is needed. All listed metrics for merging and restructuring are marked with YES. This should be interpreted as a traditional validation of the global integrated schema. In the examples in the following discussion, we therefore exclude this phase since, according to the framework, the schema under development should be evaluated based on all quality metrics.

In schema integration, it is important that each stakeholder schema fulfills the completeness quality factor within each phase. However, not all metrics are applicable in each phase and it should be noted that since we are working with intermediate schemata in the integration phase, the definition of completeness could vary depending on which phase is in focus. For instance, in pre-integration it is important that each stakeholder schema includes all requirements and nothing but the specified requirements (M1 & M2) and that intra-schema conflicts are resolved (M3). In comparison of the schemata, it is important that all similarities and differences between two schemata are recognized (M3) and resolved in conforming the schemata. If a dynamic/behavioral schema has already been developed, it can for instance in pre-integration be checked for any inconsistencies between the two schema types. However, as pointed out in [6], integration of the structural schemata should take place before integration of the behavioral schemata, making this metric unusable. The schema integration process is very complex as such and the dynamic/behavioral schema should therefore not be used in the second and third phase of the integration process to reduce the complexity.

The second quality factor is ***integrity*** and refers to if and how business rules and/or integrity constraints are represented in a correct way in the schema. Integrity is measured on the basis of two metrics: number of business rules that are not represented in the schema (M5) and number of inaccurately defined integrity constraints (M6) (see Table 2).

Table 2. Quality Factor 2 *Integrity* and the Integration Process

Metric	Pre-Integration	Comparison of the Schemata	Conforming the Schemata	Merging and Restructuring
M5	YES	YES	YES	YES
M6	YES	YES	YES	YES

In schema integration, the integrity quality factor can for instance be compared with recognition of dependency conflicts (M6). One problem that M5 might highlight is whether it is possible to specify all business rules using the chosen modeling language. For some business rules, modeling the structural part is not enough. One example of this is given in the update problem having the concepts of Product, Product.Quantity, OrderLine and OrderLine.Quantity. Suppose that we need to reduce the ordered amount for a specific product: we then need to decrease the specific OrderLine.Quantity for the ordered Product and at the same time increase the Product.Quantity value since we have more items to sell. However, there is no natural connection between Product.Quantity and OrderLine.Quantity. We therefore need to specify this in natural language text or describe the scenario using a modeling language suitable for the dynamic/behavioral part. A longer discussion and a proposed solution for the described problem are addressed in [5].

The third quality factor is ***flexibility*** and refers to how the schema could cope with future business changes. Flexibility is measured on the basis of three metrics: number of schema elements that might change (M7), estimated cost of changes (M8) and strategic importance of changes (M9). This quality factor is out of the scope of what we define as schema integration. We therefore leave flexibility for now and view it as a quality factor to include and use in relation to or after the last phase in the integration process.

The fourth quality factor is ***understandability*** and refers to how easy the schemata can be understood by the stakeholders. Understandability is measured on the basis of three metrics: how the users rate the understandability of the schema (M10), if the schema is actually understood by the users (M11) and how the application developers rate the understandability of the schema (M12) (see Table 3).

In schema integration, the stakeholders are an important source of domain knowledge and should therefore be involved during the whole integration process. Involving the stakeholders is important not only in manual integration but also in semi-automatic integration [6].

Table 3. Quality Factor 4 *Understandability* and the Integration Process

Metric	Pre-Integration	Comparison of the Schemata	Conforming the Schemata	Merging and Restructuring
M10	YES	YES	YES	YES
M11	YES	YES	YES	YES
M12	NO	NO	NO	YES

Understandability is therefore also an important quality factor and metrics M10-M11 are therefore applicable in all phases. It should be noted however that a user might *believe* that s/he understands the global integrated schema and therefore rate it high (M10). For this reason, it is important to combine metrics M10 and M11 since M11 addresses whether a schema is actually understood and not just if the schema is understandable.

The fifth quality factor is ***correctness*** and refers to whether the schema follows the rules of the chosen modeling language. Correctness is measured on the basis of three metrics: number of errors in relation to the rules of the chosen modeling language, (M13), number of errors in relation to the 1st, the 2nd, the 3rd and 4th+ normal form (M14) and the number of redundant schema elements between concepts (M15) (see Table 4).

Table 4. Quality Factor 5 *Correctness* and the Integration Process

Metric	Pre-Integration	Comparison of the Schemata	Conforming the Schemata	Merging and Restructuring
M13	YES	YES	YES	YES
M14	NO	NO	NO	YES
M15	NO	NO	NO	YES

The rules of the chosen modeling language should of course be applied (M13) since they dictate how the schemata should be constructed. However, our point of departure is that there exist schemata on different levels of abstraction implementation-independent and implementation-dependent and in this paper our focus is on implementation-independent. Therefore we do not deal with normalization, which is a task to perform in logical database design. Metric 14 is therefore not applicable in the integration process. The third metric (M15) for correctness is also marked with NO since the schemata is designed on an implementation-dependent level, where redundant concepts such as synonyms should be kept as long as possible in the integration process.

The sixth quality factor is ***simplicity*** and refers to the number of modeling elements in the schema. This quality factor is measured on the basis of three metrics: the number of concepts (M16), the number of concepts and connections (M17) and the number of schema elements (M18) (see Table 5).

Often it is possible to model the same phenomenon using different modeling patterns within one and the same modeling language (e.g. [5]) and according to the rules in the framework, the pattern and schema using the fewest modeling elements should be used. Metric 18 is marked with both a YES and a NO indicating that this specific metric is only applicable for schemata modeled using a modeling language that distinguishes between entities/classes and attributes and not for modeling languages that only focus on concepts and connections between concepts (e.g. ORM [10]).

Table 5. Quality Factor 6 *Simplicity* and the Integration Process

Metric	Pre-Integration	Comparison of the Schemata	Conforming the Schemata	Merging and Restructuring
M16	YES	YES	YES	YES
M17	YES	YES	YES	YES
M18	YES/NO	YES/NO	YES/NO	YES

The seventh quality factor is ***integration*** and refers to how consistent the schema is in relation to other data used within the organization. This quality factor is measured on the basis of four metrics: number of conflicts in relation to the 'master' organizational schema (M19), number of conflicts in relation to already implemented information systems (M20), number of data elements that are already stored in implemented information systems and projects (M21), and ratings from adjacent business areas whether the definitions of schema elements fit into the organization and not just the application being developed (M22) (see Table 6).

Table 6. Quality Factor 7 *Integration* and the Integration Process

Metric	Pre-Integration	Comparison of the Schemata	Conforming the Schemata	Merging and Restructuring
M19	YES	YES	YES	YES
M20	YES/NO	YES/NO	YES/NO	YES
M21	YES/NO	YES/NO	YES/NO	YES
M22	YES	YES	YES	YES

Different knowledge sources such as domain ontology, taxonomy and dictionary are often used to facilitate the schema integration process (e.g. [6]). The usage of such a knowledge source has similarities with the usage of the "master" organization schema and is therefore applicable in the schema integration process (M19).

In [2], the authors distinguish between schema (view) integration and database integration. The former relates to conceptual database design and the latter to design of a distributed database. Database integration results in a global schema in which all local database schemata are incorporated into "a virtual view of all databases taken together in a distributed database environment" (p. 324) [2]. Therefore M20 and M21 mostly refer to database integration, but if taking data already implemented in other information systems into account, these metrics are also applicable in the schema integration process (see YES/NO for phase 1-3).

Finally, viewing the database as being part of the organization and not as a standalone database encourages the rating from adjacent business areas (M22).

The eighth and last quality factor is ***implementability*** and refers to whether the schema can be translated during logical database design and implemented during the physical database design within the stated limitations. Implementability is measured on the basis of three metrics: rating risks in relation to the chosen technology (M23), rating risks in relation to the given schedule (M24) and estimation of development costs (M25). Because this quality factor is viewed as out of the scope of what we define as schema integration, we leave implementability for now and view it as a quality factor to include and use in relation to or after the last phase in the integration process.

6 Summary and Conclusions

In this paper we have addressed schema quality in the schema integration process. In doing so, we have combined the framework for schema quality developed and described in [20][21][22] with the schema integration process as described in [2]. In doing so we mainly focused on what is often referred to as the implementation independent level producing implementation neutral schemata. Focusing on the implementation independent level indicates that we have studied the problem of schema quality and process quality within the schema integration process from a new perspective, adding contributions to the research field of schema integration. As its main contribution, the paper not only offers new insights on how to improve the quality of the integration process but also suggests that the definition of what a high quality schema is differs between the four phases in the schema integration process.

As a next step, we plan to describe concrete tasks that should be performed to improve the schema quality within the several steps of schema integration.

References

1. Batini, C., Ceri, S., Navathe, S.B.: Conceptual Database Design an Entity Relationship Approach. Addison Wesley Publ. Comp. (1992)
2. Batini, C., Lenzerini, M., Navathe, S.B.: A Comparative Analysis of Methodologies for Database Schema Integration. ACM Computing Surveys. 18(4), 323--364 (1986)
3. Becker, J.: Die Grundsätze ordnungsmäßiger Modellierung und ihre Einbettung in ein Vorgehensmodell zur Erstellung betrieblicher Informationsmodelle. (1998) http://www.wi-inf.uni-duisburg-essen.de/MobisPortal/pages/rundbrief/pdf/Beck98.pdf (2012-06-29)
4. Bellström, P.: On the Problem of Semantic Loss in View Integration. In Barry, C., Conboy, K., Lang, M., Wojtkowski, G., Wojtkowski, W. (eds.) Information Systems Development Challenges in Practice, Theory, and Education. Volume 2, pp. 963--974. Springer, New York (2008)
5. Bellström, P.: Schema Integration How to Integrate Static and Dynamic Database Schemata. Dissertation. Karlstad University Studies 2010:13 (2010)
6. Bellström, P., Vöhringer, J.: A Semi-Automatic Method for Matching Schema Elements in the Integration of Structural Pre-Design Schemata. International Journal on Advances in Intelligent Systems. 3 & 4, 410--422 (2011)
7. Cherfi, S.S., Akoka, J., Comyn-Wattiau, I.: Perceived vs. Measured Quality of Conceptual Schemas: An Experimental Comparision. In: Grundy, J., Hartmann, S., Laender, A. H. F., Maciazek, L., Roddick, J.F.: Proceedings of Tutorials, Posters, Panels and Industrial Contributionof the Twenty-Sixth International Conference on Conceptual Modeling (ER 2007), CPRIT, vol. 83. (2007)
8. Ekenberg, L., Johannesson, P.: A Formal Basis for Dynamic Schema Integration. In Thalheim, B. (ed.) Conceptual Modeling – ER'96. LNCS 1157, pp. 211--226. Springer, Berlin (1996)

9. Frank, U.: Integration – Reflections on a Pivotal Concept for Designing and Evaluating Information Systems. In Kaschek, R., Kop, C., Steinberger, C., Fliedl, G. (eds.) Information Systems and e-Business Technologies. LNBIP 5, pp. 111--122. Springer, Berlin (2008)
10. Halpin, T., Bloesch, A.: Data modelling in UML and ORM: a comparison, Journal of Database Management, IGI Global, 10 (4), 4--13 (1999)
11. Hevner, A.: A Three Cycle View on Design Science Research. Scandinavian Journal of Information Systems. 19(2), 87--92 (2007)
12. Hevner, A., Chatterjee, S.: Design Research in Information Systems Theory and Practice. Springer, New York (2010)
13. Hevner, A., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quartely. 28(1), 75--105 (2004)
14. Iivari, J.: A Paradigmatic Analysis of Information Systems As a Design Science. Scandinavian Journal of Information Systems. 19(2), 39--64 (2007)
15. Johannesson, P.: Schema Integration, Schema Translation, and Interoperability in Federated Information Systems. Dissertation. Stockholm University & Royal Institute of Technology No. 93-010-DSV (1993)
16. Lee, M.L., Ling, T.W.; A Methodology for Structural Conflict Resolution in the Integration of Entity-Relationship Schemas. Knowledge and Information Systems. 5, 225--247 (2003)
17. Lindland, O.I., Sindre, G., Solvberg, A.: Understanding Quality in Conceptual Modeling. IEEE Software, March, IEEE Press (1994)
18. Mannino, M.V.: Database Design, Application Development, & Administration. New York: McGraw-Hill (2007)
19. March, S.T.; Smith, G.F: Design and Natural Science Research on Information Technology. Decision Support Systems. 15, 251--266 (1995)
20. Moody, D.L.: Metrics for Evaluating the Quality of Entity Relationship Models. In Ling, T.W., Ram, S., Lee, M.L. (eds.) Conceptual Modeling – ER'98. LNCS 1507, pp. 211--225. Springer, Berlin (1998)
21. Moody, D.L., Shanks, G.G.: What Makes a Good Data Model? Evaluating the Quality of Entity Relationship Models. In Loucopoulos, P. (ed.) Entity-Relationship Approach – ER'94. LNCS 881, pp. 94--111. Springer, Berlin (1994)
22. Moody, D.L., Shanks, G.G.: Improving the Quality of Data Models: Empirical Validation of a Quality Management Framework. Information Systems Journal, 28(2), 619--650 (2003)
23. Moody, D.L., Theoretical and Practical Issues in Evaluating the Quality of Conceptual Models: Current State and Future Directions. Data & Knowledge Engineering. 55 (3), 243--276 (2005)
24. Navathe, S., Elmasri, R., Larson, J.: Integrating User Views in Database Design. IEEE Computer. 19(1), 50--62 (1986)
25. Parsons, J.: Effects of Local versus Global Schema Diagrams on Verification and Communication in Conceptual Data Modeling. Journal of Management Information Systems. 19(3), 155--183 (2003)
26. Song, W.: Schema Integration – Principles, Methods, and Applications. Dissertation. Stockholm: Stockholm University & The Royal Institute of Technology No. 95-019 (1995)
27. Spaccapietra, S., Parent, C.: View Integration: a Step Forward in Solving Structural Conflicts. IEEE Transactions on Knowledge and Data Engineering. 6(2), 258--274 (1994)