

ISO/IEC JTC 1/SC 32 WG2 N1066

Last Updated: 2007-12-04

ISO/IEC JTC 1/SC 32/WG 2

Metadata

DOCUMENT TYPE	Expert contribution
TITLE	A NIAM2007 Conceptual Analysis of the ISO and OMG MOF Four Layer Metadata Architectures
SOURCE	Sjir Nijssen , PNA Group, Netherlands
STATUS	Paper presented at ORM 2007 provided to SC32 WG2 for information.
REQUESTED ACTION	For information
Number of Pages	12 plus cover page
LANGUAGE USED	English
DISTRIBUTION	For WG2 Convener and P & L Members ONLY . Not for general distribution.

A NIAM2007 Conceptual Analysis of the ISO and OMG MOF Four Layer Metadata Architectures

Inge Lemmens, Maurice Nijssen, and Sjir Nijssen

PNA Group, Geerstraat 105, 6411 NP Heerlen, The Netherlands
{inge.lemmens, maurice.nijssen, sjir.nijssen}@pna-group.nl

Abstract. OMG and ISO gain more and more influence in IT by issuing standards to the world community. The better the quality and understandability of such standards the easier they get accepted. Important standards of ISO and OMG are based on the so-called classical framework for meta-modeling. In this paper we will use the NIAM2007 conceptual schema specification procedure or modeling processes to *derive* the layers of the metadata architecture. We propose to accept the outcome of logical reasoning as a solid basis for understanding.

1 Introduction

Knowledge in the world of IT is increasing in volume as well as complexity. In the world of web applications, the complexity and size have reached a point where one gets concerned about the applicability of all the knowledge. Therefore, there is a growing need to maximize the use of explicit and declarative knowledge over procedural knowledge [7, 17]. Moreover, in such an increasingly complex world it is recommended to develop and use conceptual foundations which are solidly based on logical reasoning and are understandable to a large number of people. The Meta Object Facility (MOF) [3, 9, 13, 15, 16] is an important step in the right direction. The MOF defines a framework for specifying, constructing and managing technology neutral meta-models. “The central theme of the MOF approach to metadata management is extensibility. The aim is to provide a framework that supports any kind of metadata, and that allows new kinds to be added as required. In order to achieve this, the MOF has a layered metadata architecture that is based on the classical four layer metamodeling architecture popular within standards communities such as ISO and CDIF. The key feature of both the classical and MOF metadata architectures is the meta-metamodeling layer that ties together the metamodels and models.” [9]

In the IT community there is an increasing awareness of the need to have solid processes to develop the meta-models. “We are concerned with a core aspect of the processes of obtaining conceptual models.” [2] The NIAM community has been working since the early seventies on the development of processes for developing conceptual models and continues to refine these as more and more industrial strength practices become available [11].

We will apply the NIAM2007 processes for developing a conceptual schema to the MOF four-layer meta-modeling architecture. The NIAM2007 processes for developing a conceptual schema are the same for the domain-specific and generic conceptual schema [1, 6, 12]; said otherwise the NIAM2007 processes are the same for the domain-specific and generic level. In this paper we will start with the same example as is used in the OMG and ISO specification [9, 13] and apply hereupon the NIAM2007 conceptual schema specification processes. It is our belief that there exists a miscomprehension about the fourth layer of the popular, classical and omni-accepted meta-modeling architecture. We will show that the third layer of the architecture, is the layer in which the meta-models are already self-describing, removing the necessity for a fourth layer.

If our analysis results are essentially different from the ISO and OMG results, there are two questions to be asked:

1. Are the NIAM2007 processes right? If not, forget this paper; if they are, then
2. What to do next?

2 From the World of Concrete Examples to the Domain-Specific Conceptual Schema

In the OMG [13] and ISO [9] specifications a specific example is used to illustrate the four layer metadata architecture. We will use the same example for easy comparison. Let us assume we see part of a ticker tape as represented in figure 1.



Fig. 1. A stock ticker tape section

According to the NIAM2007 processes for developing a conceptual schema, available since a long time [11] and extensively described by Halpin in [4], the conceptual analyst asks a domain expert to verbalize the contents of the concrete example to a colleague over the telephone. In this case we would get the contents of table 1.

Table 1. Result of verbalization at the level of ground facts

The latest price of the company Sunbeam Harvesters is 98.77.
” ” ” ” ” ” ” Ace Taxi Cab Ltd ” 12.32.

According to the NIAM2007 conceptual schema specification processes, the constant parts are taken as the basis for the fact type reading [18] (also called sentential form [18], sentence pattern [12]) of the next level conceptual schema, in this case the domain-specific conceptual schema, and the variable parts are replaced by a role or placeholder. Hence for the many fact instances sharing the same constant

parts and sharing variable parts which perform the same role, we introduce a fact type reading or sentence pattern. If we apply this procedure in this concrete case we obtain:

1: The latest price of the company <Company> is <Price>.

Probably the boys in the stock office would abbreviate this to

2: <Company> <Price>.

Hence we have now obtained two sentence patterns [12] or fact type readings [18] for two different communities. In SBVR [18] one would say there is one “semantic community whose unifying characteristic is a shared understanding (perception) of the things they have to deal with” [18] (Clause 11.1.1.1) and two speech communities “whose unifying characteristic is the vocabulary that it uses” [18] (11.1.1.1). There exists a business rule of necessity here: Each speech community is of exactly one semantic community [18]. Hence, the two speech communities share the same conceptual structure. This shared conceptual structure is the fact type. According to NIAM2007, the fact type is diagrammatically represented in figure 2.

StockQuote



1: The latest price of the company <Company> is <Price>.

2: <Company> <Price>.

Fig. 2. Intermediate fact type diagram

As you can see we have, for comparison reasons, given the fact type the same name as was given to the record at the model level of the MOF Four Layer Metadata Architecture; the same holds for the roles, called fields in the model level of MOF [13].

We may conclude that there is an easy to follow conceptual schema specification process to arrive at the core of the domain-specific conceptual schema. The core is taken to include primarily the data structure. We take figure 6.1 of the MOF [13] and present additionally in that figure which road (or process) we have traveled to go from the concrete example (in MOF called the information level) to a part of the domain-specific conceptual schema (in MOF called the model level). The result is presented in figure 3 (the arrow labeled (see section 2), going from the information to the model level).

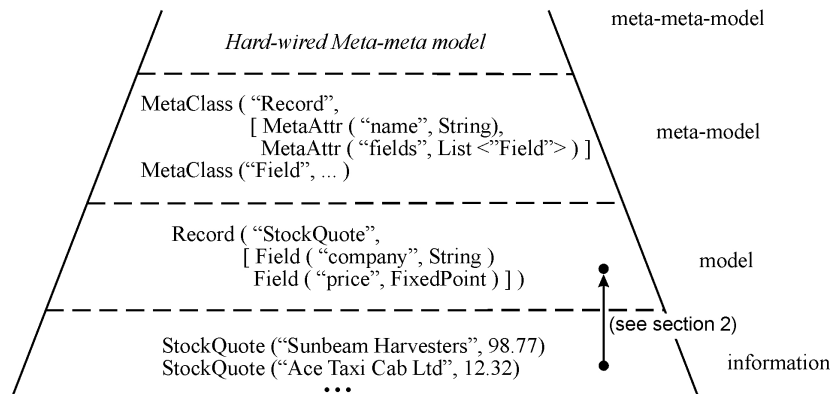
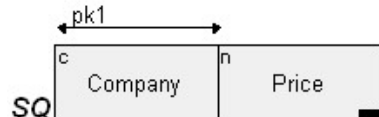


Fig. 3. The process from the world of concrete examples to the domain-specific conceptual schema, superimposed on the MOF diagram

In NIAM2007 it is strongly recommended to specify the primary uniqueness constraint for every fact type, even if there is hardly any time to do so. What is the value of a function if it is not known which are the independent and which the dependent variables! And if there is a little more time add the most relevant constraints right away. If we do this we get the contents of figure 4.

StockQuote



- 1: The latest price of the company <Company> is <Price>.
- 2: <Company> <Price>.

Fig. 4. Complete fact type diagram at the domain-specific conceptual schema level

The arrow, labeled pk1, above the role Company specifies that in the population of the facts (in MOF terms the level labeled information) conforming to this fact type, there are no duplicates. The little black rectangle in the lower right corner of role Price indicates that for every fact in this fact population the price must be known. The value of the Company is already known for every fact as the primary uniqueness constraint means two matters: no duplicates beneath me and no empty cell in any fact beneath me. If we would only use this diagrammatic convention for *elementary* fact types, then the little black rectangle is superfluous. However, we use the same diagrammatic formalism for gracefully combined fact types. The observation of

thousands of gracefully combined fact types indicates an average of 10 roles per such fact type.

Associated with this fact type diagram are two fact type readings. One is for the normal business persons or customers interested in stock prices, one for the fast boys on the floor. Hence in SBVR terms, we have here one semantic community (represented by the fact type) and two speech communities (each represented by the associated fact type reading) [18]. Both speech communities share the same meaning (in this case fact type and constraints) but prefer a different language to express the meaning (fact type readings). We have now completed part I of our journey, and will start the second leg of this journey.

3 From Domain-Specific to the Generic Conceptual Schema

In the NIAM2007 approach an information bearing construct at any level is considered a fact [12]. Hence we can take the result of the previous specification process (figure 4) and take that as input for the next conceptual schema specification process. If we do that we have the option to start with the verbalization of the fact types, roles, the primary uniqueness constraint, the mandatory constraints, or the fact type readings. In order to stay as close as possible to the illustration of the four layer metadata architecture of ISO and OMG [13] we will verbalize the roles (using type II verbalization).

This type of verbalization (type II) has as its aim to transform the information bearing constructs into facts without keywords [18], hence simple declarative facts, into a format that makes the step towards the next level of grammar easy. In this case, the information bearing constructs form the domain-specific conceptual schema rules. We call this in NIAM2007 Verbalization type II. There are no keywords such as “each”, “at least one”, “at most one”, “must”, “must not”, “always”, “never”, “another”, using SBVR terminology, in this type of verbalization. The audience of type II verbalization is quite different from the SBVR primary audience, namely business people. The two types of verbalization have the same result at the level of the facts (without any grammar function) but have quite different results at the level of any conceptual schema, be it domain-specific or generic. Verbalization type I is aimed at the business person, verbalization type II at the conceptual schema expert. Verbalization type II is quite different from the verbalization as described in Halpin [4, 5] and SBVR [18], which both describe type I. Verbalization type I is also available in NIAM2007 and has the same aims as described in Halpin and SBVR, namely the transformation of a presentation into facts. The authors are strongly convinced that in practice both types of verbalization are useful.

Also note that one fact type can serve as the basis for any number of fact type readings or sentence patterns. This kind of diagram may contain furthermore a population as well as the verbalized facts using the fact type readings.

The result of the process verbalization type II, applied to the roles in figure 4 is shown in table 2.

Table 2. Result of verbalization at the metadata (domain-specific conceptual schema) level

Role	Company	of fact type	StockQuote	takes position	1.
"	Price	"	"	"	StockQuote
"		"		"	2.

If we now apply the generalization process (as shown previously in section 2) we will produce a fact type reading, thereby arriving at the generic level, or the meta-model level in OMG MOF terminology:

1000: Role <RoleName> of fact type <FactType> takes position <Position>.

The fact type diagram generated from the fact type reading is represented in figure 5.

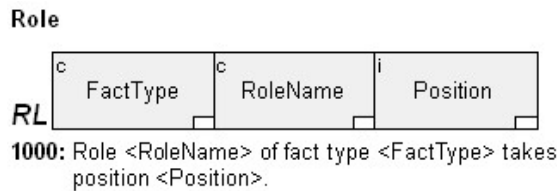


Fig. 5. Intermediate fact type diagram

There is a specification process to add the constraints but space does not permit us to describe this and therefore we present the results as in figure 6. Primary uniqueness constraint pk1001 specifies that there are no duplicates in the fact population under the combination of the roles FactType and RoleName and alternate key ak1002 specifies there are no duplicates under the combination of the roles FactType and Position. The black rectangle in the right lower part of the role Position specifies that every fact in the fact population must have a value for the role Position.

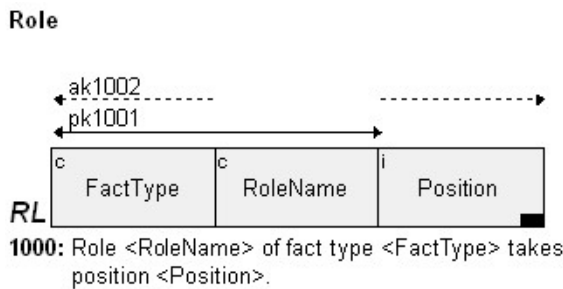


Fig. 6. Complete fact type diagram at the generic conceptual schema level

We have now finished two legs of our journey to discover by logical reasoning how many layers there should be in the MOF architecture. In the next section we will apply the same NIAM2007 conceptual schema specification processes to the fact type of figure 6 to see where we end up, at MOF level 4 or ...

4 From the Generic Conceptual Schema to the Next Level

We have applied the two NIAM2007 processes verbalization type II and generalization, to the ground fact level and the domain-specific conceptual schema level. In figure 7 these processes are added to the figure as represented in the MOF specifications.

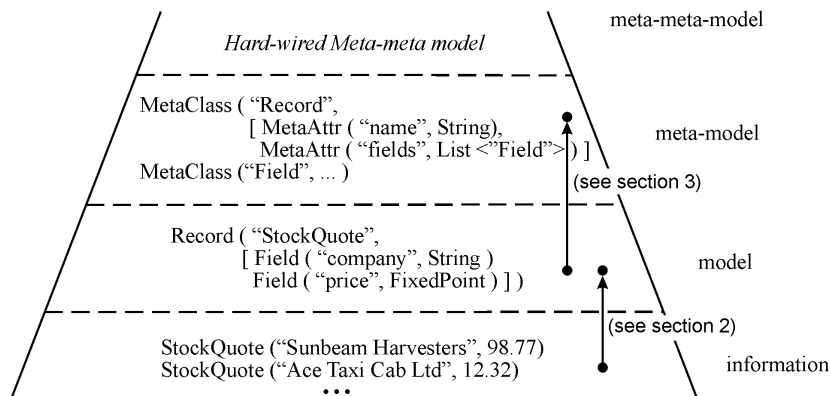


Fig. 7. The process from the domain-specific conceptual schema to the generic conceptual schema, superimposed on the MOF diagram

The question is: where will we end up if we apply *exactly the same* conceptual schema specification processes to the meta-model? This question is illustrated in figure 8.

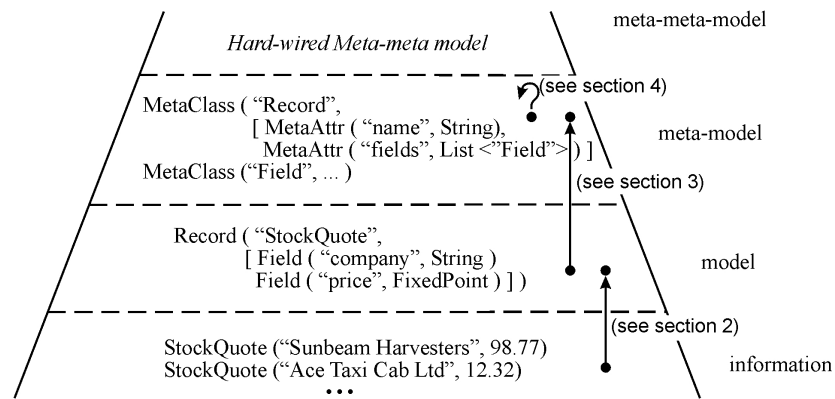


Fig. 8. Where do we arrive if we start the conceptualization process at the generic conceptual schema level?

It is an assumption of NIAM2007 that verbalization type II can be applied at any level, hence also at the meta-model level. This was already accepted as a working axiom in 1980 [11]. It is a pleasure to read in a recent article (2007): “One common characteristic of software analytics tools is that they view software as data.” [10] Let us do this for the facts about roles (of figure 6) and we get the following:

Table 3. Result of verbalization at the meta-metadata (generic conceptual schema) level

Role	FactType	of fact type	Role	takes position	1.
”	RoleName	”	”	”	2.
”	Position	”	”	”	3.

If we apply the generalization process to these three facts we obtain the associated fact type reading and we give this fact type reading a different identification from all existing identifications for fact type readings and therefore 1001.

1001: Role <RoleName> of fact type <FactType> takes position <Position>.

The fact type diagram generated from the fact type reading would be as represented in figure 9.

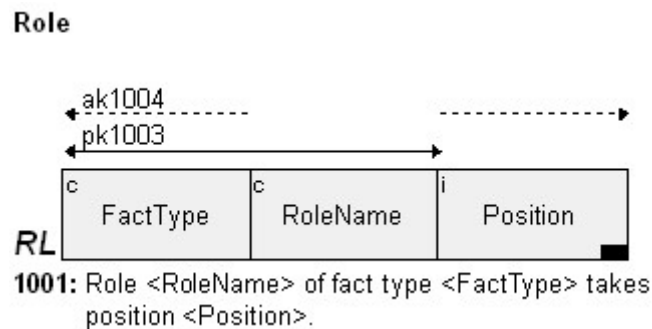


Fig. 9. Complete fact type diagram at the generic conceptual schema level

Careful analysis of the fact type of figure 6 and the fact type of figure 9 results in the conclusion that these fact types are equal and that the fact type readings 1000 and 1001 are equal.

Of course we could apply the same NIAM2007 conceptual schema specification processes to arrive at the various other meta fact types like FactType, Constraint, PrimaryKey, AlternateKey, ForeignKey, ValueRule etc. The results lead to the same conclusion as has been drawn time and again.

Conclusion: if we apply in a consistent way the NIAM2007 processes of

1. Verbalization type II and
2. Generalization toward fact type reading,

the result is that the meta-meta-model of the MOF is exactly the same as the meta-model as indicated by arrow 8 in figure 10. Arrow 8 shows that the result of the

conceptualization processes applied to the generic conceptual schema is (a proper subset of) the generic conceptual schema.

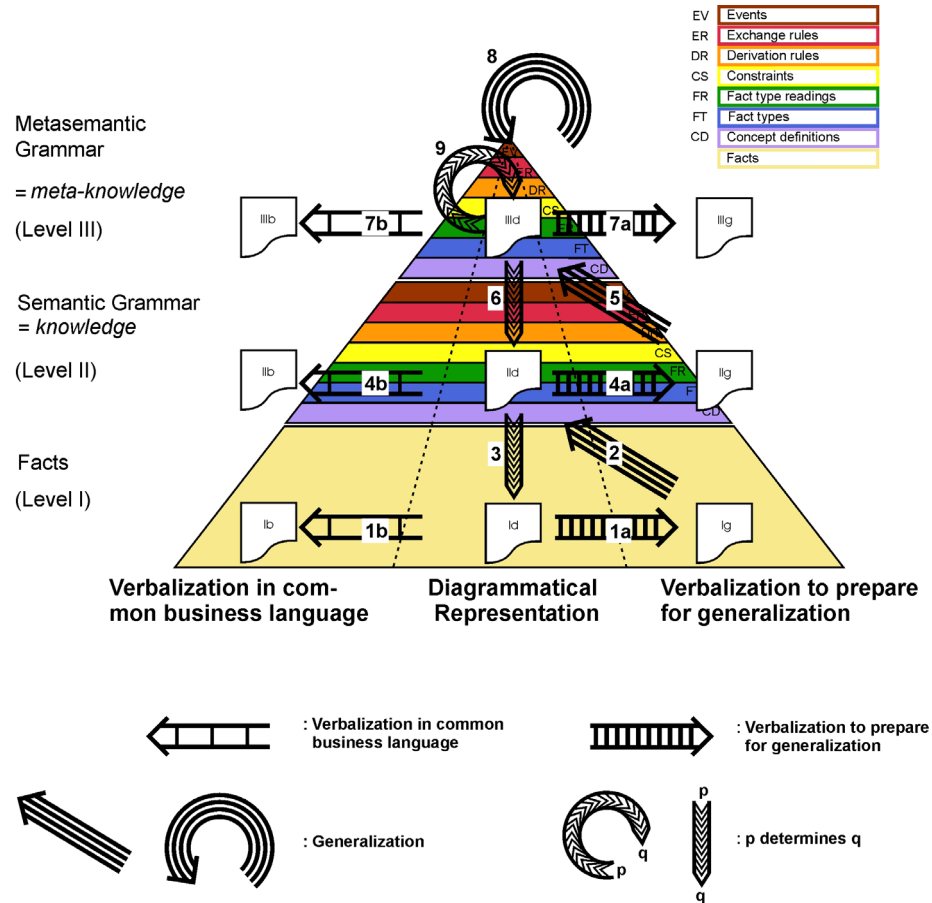


Fig. 10. Knowledge triangle with three forms of representation and two kinds of verbalization processes

The following question could legitimately be asked: what is the advantage of this OMG and ISO illogical four layer metadata architecture? The devil named inconsistency should not have a room in the clarity hotel.

5 Some Questions

As a member of the fact orientation research community we could not resist asking the following questions:

1. How could it happen that such a fundamental architecture is defined without proof?
2. Is this the only logical problem in the large specification?
3. Are there no verifiable conceptual schema specification processes used at OMG or ISO?
4. Where is the quality control in OMG and ISO?
5. Why did the fact orientation community's radar miss this target image for such a long time?
6. Is architecture of such little interest in the ORM conferences that it has not been on the agenda before?

“The key feature of both the classical and MOF metadata architectures is the meta-meta-modeling layer that ties together the meta-models and models.” [13] We hope this will soon be corrected. Using the proven conceptual schema specification procedures of NIAM2007 it was fairly straightforward not only to demonstrate how the second level (model in MOF terms; domain-specific component of the conceptual schema in SBVR terms) and the third level (meta-model in MOF terms; generic component of the conceptual schema in SBVR terms) are precisely related, but also how the first level (information in MOF terms; ground fact population in SBVR terms), second level and third level are precisely related. The view that all three levels contain declarative facts is one that was adopted by the NIAM community in the late seventies [11], and we are pleased to note also by OMG [14, 10]. Since the late seventies, considerable improvements and refinements have been introduced continuously in NIAM, mainly in response to users in business practice. This view has been gradually accepted and even software is these days viewed as data. The total coherence is easier to maintain if all information bearing constructs are viewed as facts, consistent use is made of fact type readings for communication purposes and concept definitions for understanding [12].

6 Summary and Recommendation

In this paper we have consistently applied the NIAM2007 specification processes to the information level, the model level and the meta-model level of the MOF architecture. In doing so it became obvious that the meta-meta-model of MOF is the same as the meta-model of MOF. Why then introduce a fourth ghost level in such an important specification? Our recommendation is to remove this inconsistency, as soon as practicable, from the MOF and ISO specifications in order to obtain a more solid understanding of the advantages of the MOF approach in the world at large. “The central theme of the MOF approach to metadata management is extensibility.” [13] This can much easier be obtained if the foundations are solid and do not lead to contradiction. Such contradictions could be demonstrated by students having attended

a 5-day NIAM2007 class, where considerable time is spent acquiring the competence to develop for each and every applicable situation the associated conceptual schema, be it domain-specific or generic.

7 Acknowledgements

A special word of thanks for the excellent reviews. One review states: “I like the fearless challenge to the omni-accepted 4-level architecture; and for me, at least, this paper has put in grave doubt the legitimacy of that architecture. (I am embarrassed to say that I've never seriously questioned it until now.) Although someone could perhaps come up with a coherent, intelligible distinction between level 3 and level 4, I think this paper makes it very doubtful that that distinction could be relevant enough to legitimize a division into these two levels.” Another review says: “The paper proves well that NIAM2007 can be used as a meta-meta model to define itself as a meta model, but it stops short of proving that it is capable of defining all meta models (such as: SysML, CWM, SBVR, BPMN, UML, ER) - which is the goal of the meta-meta model in the MOF 4-layer view.” In [11] (1980) one can read: “All grammars of the Coexistence information systems framework can be considered as information bases.” In current day language this means that the internal schema, be it UML or ER, can be and has been expressed by the authors in facts. BPMN and SBVR have also been expressed in facts by the authors. Space does not permit us to deliver these “derivation” proofs here.

References

1. Bakema, G., Zwart, J. & van der Lek, H. 2000, Fully Communication Oriented Information Modelling, Ten Hagen Stam, The Netherlands.
2. Bommel, P., Hoppenbrouwers, S., Proper, E. & van der Weide, T. Exploring Modelling Strategies in a Meta-modelling Context, Proceedings ORM2006.
3. Bommel, P., Hoppenbrouwers, S., Proper, E. & van der Weide, T. Giving Meaning to Enterprise Architectures: Architecture Principles with ORM and ORC, Proceedings ORM2006.
4. Halpin, T. 2001, Information Modeling and Relational Databases, Morgan Kaufmann, San Fransciso.
5. Halpin, T. & Curland, M. Automated Verbalization for ORM2, Proceedings ORM2006.
6. Hoppenbrouwers, S., Lindeman, L. & Proper, E. Capturing Modeling Processes – Towards the MoDial Modeling Laboratory, Proceedings ORM2006.
7. ISO/IEC 10027, Information Resource Dictionary System (IRDS) framework, JTC 1/SC 32, International Standards Organization, 1990.
8. ISO/IEC 15474, Case Data Interchange Format (CDIF) framework, JTC 1/SC 7 N, International Standards Organization, 1998.
9. ISO/IEC 19502, Meta Object Facility, JTC 1/SC 32, International Standards Organization, 2005.
10. KDM Analytics, Why Knowledge Discovery Metamodel?, <http://www.kdmanalytics.com>.
11. Nijssen, S. 1980. A Framework for Advanced Mass Storage Applications, Proceedings of the MEDINFO 1980, Tokyo, The world conference on medical informatics.

12. Nijssen, S. & Bijlsma, R. A Conceptual Structure of Knowledge as a Basis for Instructional Designs, Proceedings IEEE Conference on Advanced Learning, 2006.
13. OMG, Meta Object Facility (MOF) Core Specification, version 2, January 2006.
14. OMG, Knowledge Discovery Meta-Model (KDM), KDM FTF ptc/2007-03-15, 2007.
15. OMG, Model Driven Architecture (MDA), ormsc/2001-07-01, 2001.
16. OMG, Model Driven Architecture (MDA), The architecture of choice for a changing world, http://www.omg.org/mda/executive_overview.htm.
17. Ross, R. Business Rule Concepts, Getting to the Point of Knowledge, 2005, BR Solutions LLC.
18. SBVR, Semantics of Business Vocabulary and business Rules, OMG, 2007.