International Journal of Computer Applications Technology and Research
Volume 9–Issue 06, 211-216, 2020, ISSN:-2319–8656

Multidimensional and Spatio-Temporal Design Pattern for Decision-Support Systems Schema Generation

Apollinaire Bamana Batoure
University Institute of Technology
The University of Ngaoundere
Ngaoundere, Cameroon

Kolyang
Higher Teachers' Training College
The University of Maroua
Maroua, Cameroon

Abstract: Design patterns are frequently used as a generic solution in software development, as well as in information systems, because they bring several improvements. In this work, we propose a design pattern that will be used for the generation of data schemas of spatio-temporal decisional systems. This use is done within the framework of the application of the Multidimensional Canonical Partitioning (MCP) approach, used for the design of decision-making systems.

Keywords: Design Pattern, Multidimensional Canonical Partitioning, Spatio-temporal decision system, UML profile, Archetype

1. INTRODUCTION

Patterns are among the reuse techniques that have been widely used in the design of information systems. They are presented as solutions to recurring problems [1, 2, 3, 4]. They help to reduce the production costs of applications, while ensuring better quality. They are also used as models.

A design pattern is an approved solution to a recurring design problem within a defined context [5, 31]. A Multidimensional Pattern (MP) is a design meta-model, representing a complete and generic star schema, in a domain of the company's activity, built from standard real-world entities [6, 7]. Their construction is based on empirical studies. It facilitates the task of the architect or designer of the system, by providing solutions to similar situations already experienced.

Design patterns are defined as abstract (e.g. specified in UML) and concrete (e.g. implementation archetypes) forms of general use, whose adaptation to a special development context is technically straightforward and cost-effective [8]. The abstract form can also be defined in formal language as B [9].

Design patterns represent expert knowledge, in a form that can be reused by the less expert. Because of their generic nature, they are considered microarchitectures that aim to reduce complexity, promote reuse and provide a common vocabulary for designers [10].

The purpose of using multidimensional patterns is to provide a generic solution for a domain decision problem that allows the design of data stores. It must be understandable and adaptable to the needs of the decision-makers, comprehensive, well-trained and produce the expected results. It thus reduces the time and cost of design, while preparing the way for implementation. The approach is platform-independent because it avoids, as much as possible, technological aspects that are characterized by continuous and rapid evolution [8, 11].

The research activities carried out by Jamel FÉKI's team[1] allow us to understand that the problems that have encouraged the emergence of multidimensional design patterns are:

- the fact that the design always starts from scratch, even in areas where the needs are known, and the solutions are mastered;
- the lack of capitalization of the know-how of data warehousing experts;
- the lack of reusable decision support system design approaches, and;
- the absence of techniques to assist in the specification of OLAP (OnLine Analytical Processing) requirements.

There are three types of patterns:

- creative patterns, which concern the creation of classes or objects;
- structural patterns, are interested in the composition of objects or classes to realize new functionalities;
- and behavioural patterns concern class interactions and the assignment of responsibilities [10, 12, 13].

There are three types of approaches in the use of patterns, similar to those used in the design of decision-making systems [14]. These are the bottom-up, top-down and mixed approaches [1, 8, 11].

Also, three groups of patterns exist [15, 16]. These are the:

1- Decision-making information systems team, Mir@cl Laboratory, FSEG Sfax.

- analysis patterns, which are used in the definition phase of system needs or requirements;
- design patterns, that define the abstract aspects of the system;
- and implementation patterns specified in a programming language for code generation.

In the present work, we specify a bottom-up, structural design pattern to generate multidimensional and spatio-temporal data schemas.

The purpose of this article is to introduce this pattern. For it, after this introduction where we have defined and given the interest of design patterns, we will present some cases of use in decisional systems. Then, we will bring up the MCP approach, which is the application framework of this pattern. Finally, before the conclusion, we present the profile of the archetype of the proposed pattern.

2. USE OF PATTERNS IN THE DESIGN OF DECISION-SUPPORT SYSTEMS

Several works on Decision Support Systems (DSS) take into account design patterns.

In [17] documents circulating within the company are used as Real World Entities (RWE). An association between these RWE and the Multidimensional Pattern (MP) is then made. These documents are classified on the basis of an empirical study. The proposed MP is made up of elements grouped into three categories (important, recommended and optional), distinguished on the basis of a graph. The pre-instantiation and instantiation of the MP entities is done at the logical level. Operators are defined for the instantiation of the MP. After this step, the elements that should actually make up the final data store are selected. At the physical level, relational views defined from heuristics are implemented. In order to be exhaustive, the patterns are overloaded. The application of the approach is made in the commercial domain and a tool that accompanies the approach is proposed.

In [18, 19], Feki J. et al. describe the application called Conception Assistée de Magasins de données en Étoiles (CAME) which implements the approach they propose. Emphasis is placed on the automation of the conceptual phase. The approach is divided into three main steps:

- pre-construction from source data schemas;
- automatic construction of the data store schemas to extract multidimensional concepts;
- and their validation by the designer.

Dimensions, facts and hierarchies are extracted from relational sources using heuristics.

The approach presented by Y. Hachaichi et al. in [6] is also declined in three steps, namely:

- collection and standardization of RWE;
- their classification into factual and basic entities;
- and the of DSS Schema.

The application of the approach is made in the medical field. Patterns are specifically used at the logical level, where the designer adapts the multidimensional pattern to the analytical needs. A correspondence is then made with the real-world entities, in order to propose a data store schema that best meets the business requirements.

Jones M. E. and Song I. Y. in [16] propose to set up a decision-making information system by using patterns, as early as the needs analysis phase. This approach is derived from the Who, What, Where, When, How and Why Quality Principle (QQOQCP). The approach is declined in two steps. In the first step, questions relating to the six domain patterns are asked in order to determine the associated dimensions. In the second step, questions are asked to determine the attributes associated with each dimension.

Annoni's thesis [2] proposes a tool to facilitate rapid development by reusing patterns. Patterns are grouped in a catalogue that capitalizes the development approach. The proposed tool is called electronic Business Intelligence Patterns for Analysis and Design (eBIPAD) and is dedicated to system administrators and business intelligence designers. The approach takes into account the tactical and strategic needs of decision-makers and then, the data sources to represent them intuitively.

In the present work, we use a multidimensional and spatio-temporal design pattern to derive the data schema of a decision system, from the elements of the multidimensional annotation obtained by applying the first five steps of the Multidimensional Canonical Partitioning (MCP) approach. The proposed multidimensional pattern is a continuation of the decision system design approach defined in [20] and will be used within the general framework of a model-driven architecture (MDA). This architecture makes it possible to implement all the phases (conceptual, logical and physical) of decision system modelling. The proposed multidimensional pattern is spatio-temporal in order to better understand, implement and exploit the studied phenomena.

In order to better understand the position of the proposed pattern, the following section provides a brief reminder of the MCP approach for designing and implementing decision-making systems.

3. MULTIDIMENSIONAL CANONICAL PARTITIONING APPROACH (MCP)

In [20], Batoure et al. propose a supply-driven approach, called Multidimensional Canonical Partitioning (MCP). It takes into account, without distinction, Entity/Relationship (ER) data schema. From universal relations assumption, the schema is derived into a universal relation (UR). This step provides a flat schema, where all features are grouped into a single entity. This entity is then partitioned vertically, according to characteristics or attributes semantics. To achieve it, we use a heuristic greedy type algorithm. Resulting partitions are candidates for being dimensions in the future data schema. To do this, we use an algorithm that matches the attributes present in the partitions and those that must actually be in the dimensions. Obtained dimensions are, if necessary, snow-flaked (normalized), using the third normal form algorithm (3NF). Because all the multidimensional elements are obtained, the data schema is generated using model transformations from QVT (Query View Transformation) language and a multidimensional and spatio-temporal design pattern.

The approach is recapped into six steps:

1. Verification of provided ER schema. If it is on universal relation form, we go straight to step 2, otherwise we restructure it according to universal relation assumption;

2. Vertical partitioning by fragmentation and distribution of attributes in obtained partitions;

3. Transformation of partitions into dimensions;

4. Normalization of dimensions;

5. Construction of facts table;

6. Generation of multidimensional schema.

The proposed multidimensional pattern is used in step six as a model, to transform by QVT, multidimensional elements to a data warehouse schema.

In Figure 1 and in step six of the approach, the elements of the multidimensional annotation, the design pattern and the QVT transformations are put together to obtain a specific decision data schema. The model transformations will be the subject of another work.

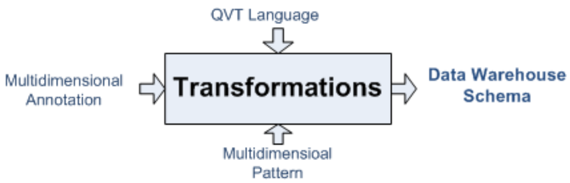

Figure. 1 Design Pattern for Approach Stage Six

We are now going to present the proposed pattern.

4. MULTIDIMENSIONAL AND SPATIO-TEMPORAL DESIGN PATTERN

The goal in this work is to propose a multidimensional and spatio-temporal design pattern, to generate data schemas for decision systems. In order to better exploit this design pattern, it is defined in a formal way by a profile, and in a concrete way by an archetype.

4.1 Design Pattern Profile

For this, we use the Unify Modelling Language (UML), which is one of the Object Management Group (OMG) standards [21]. UML enables the definition of the structural and behavioural aspects of a system via a set of models. These aspects explain the structure of a system thanks to attributes, and their behaviours thanks to operations, through classes [11, 22, 23]. The use of these models reduces the effort of the designer in learning new models and methods. Indeed, class models are unsuitable for multidimensional modelling. To overcome this deficiency and to make the semantics of multidimensional concepts explicit, we propose a UML profile.

A profile is a set of mechanisms and techniques for adapting the UML modelling language to a specific application domain. From a technical point of view, a profile is a set of stereotypes. A stereotype is defined as a domain-specific concept [24, 25]. The profile thus enables the extension of the UML meta-model. A meta-model defines a domain-specific modelling language and a profile defines a dialect or a variant of this language [15]. UML profiles involve the central concept of stereotype. It is described by a meta-class or package and several stereotypes.

In the proposed profile, stereotypes are used to define the overall data pattern, facts and their measures, then dimensions and their hierarchies. A measure in the fact table can be temporal, spatial, spatio-temporal or thematic. On the other hand, a dimension can be spatial, temporal or thematic. The dimension is described by hierarchies that follow the same characteristic as itself. The term thematic is used to describe all objects other than temporal and spatial objects [27, 28]. The spatio-temporal profile in Figure 2 has been designed using the Software Engineering Workshop (SEW) Perceptory pictograph language, also known as the "repository of user perceptions" [28, 29]. This SEW follows the UML formalism.

The proposed meta-model includes, in addition to the meta-class that defines the specific language package, four main stereotypes. These are the stereotypes on:

- the overall data schema of the decision-making system;
- the facts table (central table of the data schema) which groups together the different measures. These measures are the elements of decision-making analysis. To this main table are linked the dimension tables;
- the measures are the attributes of the fact table. They are either thematic, spatial, temporal or spatio-temporal;
- the dimensions are the tables connected to the fact table. They are either thematic, spatial or temporal. The dimensions can have granularities or hierarchies in order to give more details on the entities they describe. In this case, the global data schema is of the snowflake type. These granularities are of the same type as the dimensions they describe.

The spatial and temporal dimensions, as well as their granularity, are automatically added because the phenomena studied are defined in time and space references.

The abstract form of the design pattern allows its concrete form to be derived.

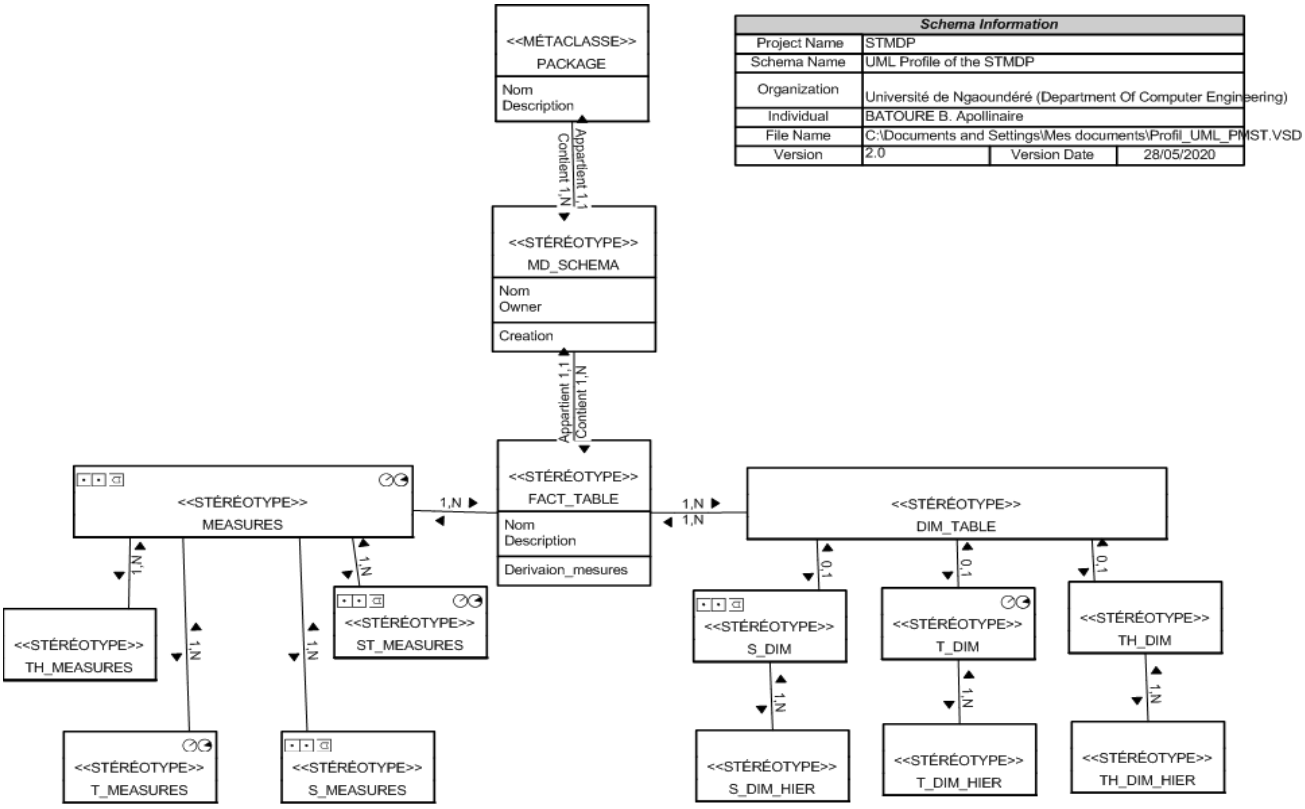

Figure. 2 Design Pattern Profile

4.2 Pattern Archetype

The concrete shape of the profile of the multidimensional design pattern allows us to define an archetype (Figure 3). It is a generic pattern obtained from the meta-model. This makes the definition of the multidimensional pattern concrete as being a standard solution, in a given domain [6, 30].

This archetype integrates the temporal and spatial dimensions as well as their granularities. It also includes thematic dimensions, the facts table and measures. The archetype can therefore be instantiated according to a specific problem from the real world. This instantiation is based on the elements of multidimensional annotation obtained in the previous steps of the MCP approach, and on models' transformation.

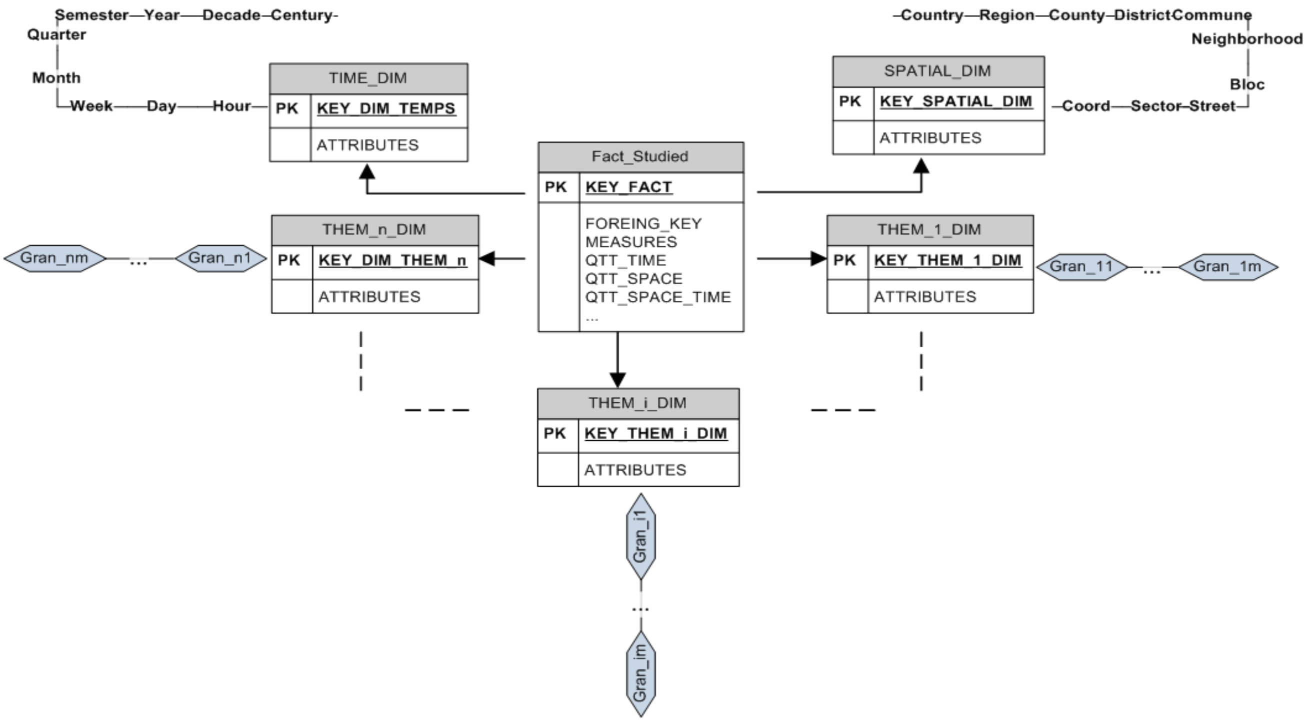

Figure. 3 Generic Data Scheme or Archetype

5. CONCLUSION

In this work, we have presented a multidimensional pattern needed for decision systems design. After defining and giving uses cases of patterns in the design of decision systems, we made a reminder of the MCP approach, which is the framework for the use of this proposal. Finally, we presented the abstract form (UML profile) and the concrete form (archetype) of this design pattern. This result, which is a continuation of the MCP decision systems design approach previously proposed and will be used within the general framework of a model-driven architecture. This architecture will enable the implementation of all modelling phases (conceptual, logical and physical) of decision system.

6. REFERENCES

[1] Lacob, M.-E., Steen, M. W. A. and Heerink, L. 2008. Reusable Model Transformation Patterns, in: Enterprise Distributed Object Computing Workshop, IEEE Computer Society, Munich, Germany, pp. 1–10.

[2] Annoni E. 2007. Éléments méthodologiques pour le développement des systèmes décisionnels dans un contexte de réutilisation, Ph.D. thesis in Computer Science, Université Paul Sabatier, Toulouse III.

[3] Gamma, E., Helm, R., Johnson, J. and Vlissides, J. 2015. Design Patterns : Element of Reusable Object Oriented Software. Addison-Wesley.

[4] Zimmer, W. 1995. Relationships between Design Patterns. In Design, volume 1, pages 345–364. ACM Press/Addison-Wesley Publishing.

[5] Miralles, A. 2006. Model Engineering for Environmental Applications, Thesis for the PhD in Computer Science at the University of Montpellier II.

[6] Hachaichi, Y., Ben-Abdallah, H. and Feki, J. 2008. Towards an MDA approach to data warehouse development, Information Systems and Business Intelligence, 158-170pp.

[7] Feki J. and Hachaichi Y. 2007. MD Assisted Design: An approach and a tool. Journal of Decision Systems vol 16/3, 303- 333.

[8] Cho, H. and Gray, J. 2011. Design Patterns for Metamodels, in: SPLASH '11 DSM Workshop, Portland, OR, USA, pp. 25–32;

[9] Marcano-Kamenoff R., Levy N. and Losavio F. 2000. " Spécification et spécialisation de patterns en UML et B " Langages et modèles à objets, LMO'2000

[10] El Boussaidi, G. 2009. Développement logiciel par transformation de modèles, Thèse en vue de l'obtention de grade de Philosophiæ Doctor (Ph.D.) en Informatique.

[11] Lano, K., and Kolahdouz R. S. 2014. Model-Transformation Design Patterns, IEEE Transactions on Software Engineering 40 (12) 1224–1259.

[12] Ergin, H., Syriani, E. and Gray, J. 2016. Design Pattern Oriented Development of Model Transformations, Computer Language, Systems and Structures.

[13] Blazy, S., Gervais, F. and Laleau, R. 2003. Une démarche outillée pour spécifier formellement des patrons de conception réutilisable, Workshop OCM-SI.

[14] Batoure, A. B. and Kolyang. 2018. Designing Decision Support Systems Approaches using Entity/Relationship Data Schema : A Survey. International Journal of Scientific & Engineering Research (IJSER), Volume 9, Issue 9.

[15] Sabri, A. and Kjiri, L. 2012. Patterns to analyze requirements of a Decisional Information System, Workshop AIde à la Décision à tous les Étages, AIDE, 12th International Francophone Conference on Knowledge Extraction and Management EGC Bordeaux.

[16] Jones, M. E. and Song, I.-Y. 2005. Dimensional modeling: identifying, classifying & applying patterns. In Song and Trujillo. 29–38.

[17] Feki, J., Ben-Abadallah, H. and Ben Abdallah, M. 2006. Reuse of star patterns. XXIVth INFORSID'06 Congress, pp. 687-701, Hammamet, Tunisia.

[18] Feki, J. and Hachaici, Y. 2007. MD Assisted Design. Journal of Decision Systems, Vol 16 - N° 03, pp 303-333.

[19] Ben Abdallah M., Feki J. and Ben-Abdallah H. 2006. MPI-EDITOR: a tool for specifying OLAP needs by logical reuse of multidimensional patterns. Decision-making Systems Workshop (ASD'06), Morocco.

[20] Batoure, A. B., Kolyang, and Tchotsoua, M. 2017. Using Multidimensional Canonical Partitioning (MCP) as a Supply-Driven Approach for Data Warehouses Design. International Journal of Computer (IJC), 25(1):52–62.

[21] Boulil, K. 2012. Une Approche Automatisée basée sur des Contraintes d'Intégrité définies en UML et OCL pour la Vérification de la Cohérence Logique dans les Systèmes SOLAP - Applications dans le domaine agri-environnemental. Thèse soutenue en vue de l'obtention de l'titre de Docteur en informatique de l'Université de Blaise Pascal - Clermont-Ferrand II.

[22] Atigui, F., Ravat, F., Teste, O. and Zurfluh, G. 2010. Démarche dirigée par les modèles pour la conception d'entrepôts de données multidimensionnelles. In IRIT (SIG/ED).

[23] Laplanche, F. 2002. Conception de projet SIG avec UML. Bulletin de la Société géographique de Liège, 42 :19–25.

[24] Kadima. H. 2005. MDA : Conception Orientée Objet Guidée Par Les Modèles. Dunud, Paris.

[25] Zubcoff J. and J. Trujillo. 2007. A UML 2.0 profile to design Association Rule mining models in the multidimensional conceptual modeling of data warehouses. Data and Knowledge Engineering, 63(1) :44–62.

[26] Aissi S. and Gouider, M. S. 2012. Spatial and Spatio-Temporal Multidimensional Data Modelling : A survey. Internatioanal Journal of Advanced Research in Computer Science and Software (IJARCCE), 1(1) :1–8.

[27] Malinowski E. and Zimanyi E. 2004. Representing spatiality in a conceptual multidimensional model. In Proceedings of the 12th annual ACM international workshop on Geographic information systems - GIS '04.

[28] Khalfi, B., De Runz, C., Faiz, S., and Akdag, H. 2015. Modélisation conceptuelle d'objets géographiques imprécis et multiples : une approche basée F-Perceptory. In Spatial Analysis and GEOmatics conference, SAGEO, pages 1–15.

[29] Zoghlami, A., De Runz, C., and Akdag, H. 2015. F-Perceptory : an approach for handling fuzziness of spatiotemporal data in geographical databases. International Journal of Spatial, Temporal and Multimedia Information Systems (IJSTMIS), 1(1) :30–62, 2015.

[30] Ben-Abdallah, H., Feki, J., and Ben-Abdallah, M. 2009. A multidimensional pattern based approach for the design of data marts', Progressive Methods in Data Warehousing and Business Intelligence: Concepts and Competitive Analytics, 23 January, pp.172–192.

[31] Sánchez Cuadrado J., Guerra E., and de Lara J. 2011. Generic Model Transformations: Write Once, Reuse Everywhere. In: Cabot J., Visser E. (eds) Theory and Practice of Model Transformations. ICMT 2011.