Scientific Research

Multidimensional Design Paradigms for Data Warehouses: A Systematic Mapping Study

Ania Cravero, Samuel Sepúlveda

System Engineering Department, La Frontera University, Temuco, Chile.
Email: ania.cravero@ufrontera.cl, samuel.sepulveda@ufrontera.cl

ABSTRACT

Data warehouses (DW) must integrate information from the different areas and sources of an organization in order to extract knowledge relevant to decision-making. The DW development is not an easy task, which is why various design approaches have been put forward. These approaches can be classified in three different paradigms according to the origin of the information requirements: supply-driven, demand-driven, and hybrids of these. This article compares the methodologies for the multidimensional design of DW through a systematic mapping as research methodology. The study is presented for each paradigm, the main characteristics of the methodologies, their notations and problem areas exhibited in each one of them. The results indicate that there is no follow-up to the complete process of implementing a DW in either an academic or industrial environment; however, there is also no evidence that the attempt is made to address the design and development of a DW by applying and comparing different methodologies existing in the field.

KEYWORDS

1. Introduction

Data warehouses (DW) are a collection of an organization's historical data of any kind. The historical data are analyzed by the decision-makers by converting the data into strategic information in order to support the decision-making process [1]. These DWs integrate a huge amount of data coming from heterogeneous data sources into a multidimensional design (MD). This model enables the users to access the data in a more natural way, by means of its structure, composed of facts (analysis measures) and dimensions (context of the factual analysis) [2]. The information stored in the facts usually represents measurements for business processes (for example, how many products are sold? How many patients are treated? How long does a given process take? etc.), and the dimensions represent the framework for analyzing these measurements (for example, time, customer or product).

The development of a DW is not an easy task, raising some difficulties such as misalignment with the business strategy and therefore failure upon implementation [3,4]. As a result, a lot of effort has been made to develop methodologies and approaches that enable the correct creation of a MD of a DW [5].

According to Winter and Strauch [6], the methodologies or approaches of the MD can be classified according to the way in which the DW requirements are obtained. These are the approaches demand-driven, supply-driven, and the hybrid approaches that seek to combine the first two.

Given the importance of DW nowadays, this article provides a comparative study of the methodologies for the MD of DW through a systematic mapping of works on the topic. The study presents the main characteristics of the activities developed in the methodologies, as well as the notations and problem areas that each paradigm contains. It is with this motivation that this study arose from our work to compile, map and summarize the primary studies on methodologies for the MD in DW. The

results indicate that there is no follow-up to the complete process of implementing a DW in either an academic or industrial environment; however, there is also no evidence that the attempt is made to address the design and development of a DW by applying and comparing different methodologies existing in the field. Finally, it is noted that the proposals that contribute with tools do so only at the prototype level.

The systematic mapping of studies is a methodology used frequently in medical research, which has been adapted for use in the IT area [7].

This work is organized as follows: Section 2 presents the problem of definition. Section 3 presents the basic concepts and a brief description of the paradigms for the MD of a DW. Section 4 describes the process of systematic mapping. Section 5 describes the results. Section 6 includes work related. Finally, Section 6 presents the conclusions.

2. Problem Definition

Although there are a variety of methodologies and approaches for the design of DW, the researchers believe that this research area is very poor. In this sense Rizzi and others authors indicate that "A very few comprehensive design methods that have been devised so far [5,8]. Overall, believe that some specific issues in design, has not been properly investigated yet. Besides, more generally, mechanisms should appear to coordinate all DW design phases allowing the analysis, control, and traceability of data and metadata along the project life-cycle" [9]. As yet there is no common strategy for the development of data warehouses [10].

On the other hand, the proposed methodologies are not always coupled with an appropriate technique for requirement analysis to form a methodological approach ensuring that the resulting database will be well-documented and will fully satisfy the user requirements [11]. In this sense, the DW is acknowledged as one of the most complex information system modules and its design and maintenance is characterized by several complexity factors that determined, in the early stages of this discipline, a high percentage of real project failures [12,13].

The awareness of the critical nature of the problems and the experience accumulated by practitioners deter-

mined the development of different design methodologies and the adoption of proper life-cycles that can increase the probability of completing the project and fulfil the user requirements [11]. For these reasons, it is suggested a survey of methodologies for DW design in order to help the reader make crucial choices more consciously.

3. Basic Concepts

3.1. Data Warehouse

The classic definition of DW was proposed by Inmon [14] as a subject-oriented, non-volatile, integrated, and time variant collection of data in support of management's decisions.

From the functional point of view, the implementation of a DW is comprised of 3 stages: (1) data extraction from different sources, (2) consistent data transformation and loading into the DW, (3) and efficient and flexible access to the integrated data using tools for end users [14].

From the development point of view, the stages consist of: (1) requirements analysis, (2) the conceptual design of the DW, (3) the logical design, (4) the physical design, (5) and the implementation via data ETL (Extraction, Transformation and Loading) [2].

The main contribution of a DW is its ability to convert data into strategic intelligence, supporting decision-making at the highest levels of an organization. This ability is supported by the OLAP tool [15], which provides end users with configurable views of data from different angles and at different aggregation levels [16].

In order to achieve OLAP consultations quickly and flexibly, the data are organized multidimensionally (known as a star schema), where the information is classified according to the facts and dimensions [2]. The facts are the numeric data or the data that represent a specific industrial activity to be analyzed. The dimensions are the individual perspectives of the data that determine the granularity (data at the detail level) adopted for the representation of a fact. The units of the facts and their values are called measurements [2]. **Figure 1** illustrates the complete process.

The methodologies for the MD analyzed establish activities for the conceptual and logical design, which are
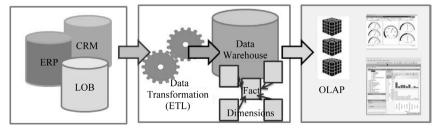

Figure 1. Process for extracting information from a DW.

classified on the basis of three paradigms that are detailed next [6].

- Supply-driven paradigm: The supply-driven approaches (also known as data-driven) initiate the DW modeling process from a detailed analysis of the data sources to determine which elements (such as facts, dimensions) are most relevant to the decision-making process.
- Demand-driven paradigm: These approaches, also known as requirement-driven or targeted, begin by determining the user's needs, then a MD of the DW is created according to the selected goals.
- Hybrid paradigm: These approaches seek to combine both paradigms in order to design the DW from the data sources, but also taking end users' needs into account. The main characteristic and difference compared to the two previous approaches is that this type can intersperse the supply- and demand-driven approaches in order to apply them at each stage of the DW development, benefitting from the information collected throughout the process.

3.2. Systematic Mapping

Systematic mapping is a reporting process and structure that can categorize the results published to date in a certain area.

The aim of systematic mapping is classification, and it is therefore directed towards the thematic analysis and identification of the main publication forums [17]. The same article indicates that it enables responses to generic questions like: What has been done to date in field X? As a limitation, this type of study does not consider the quality of the studies included.

The systematic mapping process consists of the following stages: (1) definition of the research questions, (2) scope of review, (3) execution of the search, (4) selection of the studies, (5) filtering of the studies, (6) classification scheme, (7) extraction of data and mapping processes, (8) and systematic map [17].

4. Systematic Mapping of MD Design Paradigms for DW

The primary aim of the systematic mapping of studies is to obtain an overall view of the research into the paradigms for the MD in DW. This not only entails identifying the main approaches in this area, but also their strengths and weaknesses and, of course, the future work that may take place to demonstrate possible weaknesses.

Next we describe the stages carried out.

4.1. Definition of the Research Questions

The following research questions (RQ) were defined according to the technique indicated in [18], fulfilling the proposed aims:

- (RQ1) Which paradigm do the investigations selected use most and how has the trend changed over time? This allows us to ascertain which trend this field presents, which approaches are effective and which are not.
- (RQ2) Which environment, academia or industry, is the most common when applying the research? This can help to explain the preferred environment for designing a DW.
- (RQ3) What is the contribution of the research works to the field? This can identify the contribution of the works, whether these are methodologies or approaches, and if they include tools.
- (RQ4) Which stage of the design of a DW is investigated? This can explain which stage has been researched the most: the conceptual, logical or physical design.

4.2. Scope of the Review

According to [18], the scope is defined on the basis of the following parameters.

- Population: Group of articles that describes the studies into the MD of a DW in academia and industry.
- Intervention: Any study that contains methods, approaches or tools; based on the paradigms for DW.
- Study design: Experiments, case studies, accounts of experience, research-action.
- Results: Amount and type of evidence regarding the MD of DW.

4.3. Execution of the Search

The search string consisted of Boolean expressions formed by the following key words: "data warehouse", "data warehousing", "multidimensional design", "approach", "methodology". Some of the terms were broken into boolean expressions using the OR and AND connectors, creating the following search string: ("data warehouse" OR data warehousing) AND "multidimensional design" AND (approach OR methodology).

In terms of time lapse, the search is concentrated between 1998 and 2013. This choice was made because as of 1998 several researchers began delving into this subject, using the works of Bill Immon and Ralph Kimball (considered as the "fathers" of DW), as the basis of their investigations.

The sources where the search was applied were: IEEE Digital Library, ACM Digital Library, ScienceDirect and SpringerLink.

4.4. Selection and Filtering of the Studies

In order to select the research works, we first used the inclusion criteria to analyze the title, abstract and key words, thereby obtaining the highest number of works

that make significant contributions regarding the paradigms for the MD of DW. Second, we used the exclusion criterion, where we concentrated mainly on the summary, introduction and conclusions, analyzing those works a little more where it was needed to ensure that they were relevant for the field of study.

- Inclusion criteria: books, documents, articles, theses, research works, journal and conference publications that describe the MD of a DW and that contain approaches, methodologies and/or tools in any of the following stages: conceptual, logical and physical multidimensional design.
- Exclusion criteria: (1) Works that deal with DW, but are not related to their MD. (e.g. Experiences of using a DW in the industry or in academia, data analysis with DW, business intelligence, data mining, OLAP, etc.) (2) Works that concentrate on the design of a DW, but that do not express a methodology for it. **Table 1** presents the number of articles according to the search string and the filter.

The selection process consists of three iterations performed by four reviewers. In the first iteration, each reviewer applied the inclusion and exclusion criteria to the title, abstract and key words for 10 works selected at random. A reliability of 79% was obtained according to Fleiss' kappa proposed by [19], which is very good. In the following iteration, each reviewer applied the same criteria to a set of articles that was assigned to him, now including the introduction and conclusion. In the third iteration the works where questions arose were analyzed thoroughly. This way a total of 25 relevant works were obtained for the mapping. **Table 2** presents the authors, title, year and publication source for each of these articles.

4.5. Definition of the Classification Scheme

Once the relevant articles had been selected, four types of classifications were defined based of the study objectives (see **Figure 2**):

- Paradigm developed: Model on which the articles are based, *i.e.*, supply-driven approach, demand-driven approach, and hybrid approach.
- Type of contribution: The contribution that investigation makes to the field, *i.e.*, if it is an approach, or a methodology, and if it contains a tool, or if it is me-

thodology-tool. Approach for those publications that offer new ideas or methodologies different from those already established. Methodology includes descriptions and procedures to follow and be able to carry out the MD of a DW. Tool refers to any kind of tool that helps in the process of designing a DW. Finally, methodology-tool to classify those publications that contribute with tools to apply the methodology presented.

- Application environment: The area where the investigation is developed, or where the authors aim to apply their research. In this case we classified them in two categories: academic and industry. Academic: this category is for those publications that direct their efforts to conducting new research and/or developing new ideas. Industry: this classification corresponds to those works that apply their research in some organization (for-profit or not-for-profit).
- Design stage: Development stage of a DW in which the authors concentrate on developing their research, *i.e.*, conceptual design, logical design and the physical.

4.6. Data Extraction and Systematic Mapping

After defining the classification system, the last step of systematic mapping consists of data extraction and the process of mapping the different dimensions. The complete result of this activity is in the following section. The synthesized result of our study can be seen in the bubble diagram in **Figure 3**.

Figure 3 basically illustrates two scatter plots with bubbles at the category intersections, which encompass several categories at once and give a quick, overall view of a field of study, providing a visual map. In this visualization of the results, the size of a bubble is proportional to the number of articles that are in the pair of categories that correspond to the bubble of the coordinates.

So, for example, we found 4 articles that describe a DW design methodology on the basis of the supply paradigm. Thus also, we found 7 articles that describe a methodology for the design of a DW based on the hybrid paradigm and which includes only the conceptual modeling of the DW.

We observe in the same figure there were no articles describing a methodology to develop a DW that included the stage of physical design, but there were articles that included the two first, conceptual and logical design.

Figure 4 illustrates the distribution of the works selected using a timeline. For all the publications included in the mapping (25), 11 are articles from conferences, 9 from journals, 4 are books and 1 is a thesis.

The figure provides a comprehensive framework to understand better the current state of the paradigms for the DW design and its evolution. We believe that the

Table 1. Results of the search and filter.

Search engine	IEEE	ACM	ScienceDirect	Springer	Total
Search results	89	120	51	296	701
Candidate works	30	35	16	38	173
Total relevant works	12	7	2	4	25

Table 2. Selected articles.

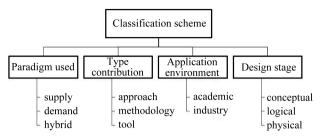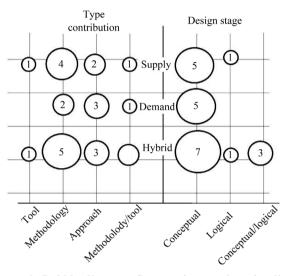| Authors	Year	Title of article	Journal or Conference
M. Golfarelli D. Maio S. Rizzi	1998	Conceptual design of data warehouses from E/R schemes	Proceedings of the Thirty-First Hawaii International Conference, System Sciences.
D. L. Moody M. Kortink	2000	From enterprise models to dimensional models: a methodology for data warehouse and data mart design	Proceedings of 2nd International Workshop on Design and Management of Data Warehouses
C. Phipps K.C. Davis	2002	Automating data warehouse conceptual schema design and evaluation	Proceedings of 4th International Workshop on Design and Management of Data Warehouses
R. Kimball L. Reeves W. Thornthwaite M. Ross	1998	The data warehouse lifecycle toolkit: expert methods for designing, developing, and deploying data warehouses	John Wiley & Sons Journal
M. Banek Z. Skocir B. Vrdoljak	2005	Logical design of data warehouses from xml	ConTEL Book
JM. Jensen T. Holmgren T. Pedersen	2004	Discovering multidimensional structure in relational data	Data Warehousing and Knowledge Discovery Journal
B. Vrdoljak M. Banek S. Rizzi	2003	Designing Web Warehouses from XML Schemas	Proceedings of 5th International Conference on Data Warehousing and Knowledge Discovery
P. Hernández I. Garrigos J.-N. Mazón	2010	Model-driven development of multidimensional models from web log files	Advances in Conceptual Modeling--Applications and Challenges In Springer
B. Hüsemann J. Lechtenbörger G. Vossen	2000	Conceptual Data Warehouse Design	Proceedings of 2nd International Workshop on Design and Management of Data Warehouses
R. Kimball M. Ross	2002	The Data Warehouse Toolkit 2nd edition	John Wiley & Sons Book
R. Winter B. Strauch	2003	A Method for Demand-driven Information Requirements Analysis in Data Warehousing	System Sciences Journal
P. Giorgini S. Rizzi M. Garzetti	2008	GRAnD: A goal-oriented approach to requirement analysis in data warehouses	Decision Support Systems Journal
N. Prakash A. Gosain	2008	An approach to engineering the requirements of data warehouses.	Requirements Engineering Journal Springer
A. Cravero J.-N. Mazón J.Trujillo	2013	A business-oriented approach to data warehouse development.	Ingeniería e Investigación Journal
L. Cabibbo R. Torlone	1998	A Logical Approach to Multidimensional Databases.	Proceedings of 6th International Conference on Extending Database Technology
M. Böhnlein A. Ulbrich-vom	2000	Business Process Oriented Development of Data Warehouse Structure.	Proceedings of Data Warehousing
A. Bonifati F. Cattaneo S. Ceri A. Fuggetta F. Paraboschi	2001	Designing Data Marts for Data Warehouses.	ACM Transactions on Software Engineering and Methodology
O. Romero A. Abelló	2006	Multidimensional Design by Examples.	Proceedings of 8th International Conference on Data Warehousing and Knowledge Discovery
O. Romero A. Abelló	2007	Automating Multidimensional Design from Ontologies.	Proceedings of the ACM tenth international workshop on Data warehousing and OLAP
JN. Mazón J. Trujillo	2008	An MDA approach for the development of data warehouses.	Decision Support Systems Journal
O. Glorio J. Trujillo	2008	An MDA approach for the development of spatial data warehouses.	Data Warehousing and Knowledge Discovery Journal
J. N. Mazón J. Lechtenbórger J. Trujillo	2011	A model-driven approach for enforcing summarizability in multidimensional modeling.	Advances in Conceptual Modeling. Recent Developments and New Directions Springer
F. Di-Tria E. Lefons F. Tangorra	2012	Hybrid methodology for data warehouse conceptual design by UML schemas	Information and Software Technology Journal
L. Gómez R. Moreno R. Pérez	2013	Computer - Assisted generation of data warehouse model: analysys of information.	DYNA Journal
M. Thenmozhi K. Vivekanandan	2012	An ontology based Hybrid Approach to Derive Multidimensional Schema for Data Warehouse.	International Journal of Computer Applications Journal

Figure 2. Classification scheme.

Figure 3. Bubble diagram. Systematic mapping visualization.

trend in the design of DW will be using hybrid approaches that consider the model of the strategy, the use of objective models, transformations between conceptual, logical and physical models, to achieve automation.

5. Comparative Analysis and Discussion

Next, from the results we answer the research questions formulated in Section 3.

- (RQ1): There are 8 articles selected according to the supply-based paradigm, 6 according to the demand-driven paradigm and 11 from the hybrid. The results indicate that most of the current research has been aimed at the hybrid approach with 11 publications corresponding to 44% of the total. One possible justification is that the authors preferred to use methodologies that reduce the failure of the DW, since this: (1) must be aligned to the organizational strategy, and (2) must have existing data fed into its data bases [4,5,13], a situation that can be addressed with this type of paradigm. The second paradigm with the greatest presence is the supply-driven approach with a total of 8 works (32%). An important point to mention is that the supply-driven approach has been neglected since 2004

due to the number of failed projects (close to 80% [13]); nevertheless, in 2010 it was taken up again, but focusing on data sources different from the organizational ones. Some examples are: data from the Web and XML schemas [20]. The paradigm with the lowest presence is the demand-driven approach with 6 articles (24%). It should be emphasized that in both paradigms, demand-driven and hybrid, there is a variety of articles in which goal models are used to represent users' needs according to the business strategy, for example i* [21,22], KAOS [23], GQM [24] and others; but only one article incorporates a process to validate the alignment between the DW and the organizational strategy [25], which is done using the standard BMM (Business Motivation Model) of the Object Management Group (OMG) [26].

- (RQ2): In terms of the environment where most of the contributions are applied (academia or industry), the numbers indicate that the authors prefer to apply their research in the industrial environment, with 15 publications selected. By contrast, we found 10 publications in the academic environment. It is worthy of note that there were no articles related to any experiment with methodologies applied to real cases, they only present a set of stages and guidelines to carry out them.

- (RQ3): In terms of the contribution of the research works (methodology, approach, tool), we can see that the methodologies were the predominant contribution in the current investigations with a total of 11 works. Second were the approaches with a total of 8 publications, and third was a combination of methodology and tool with a total of 4 publications. This type of combined contribution results from the authors presenting methodologies in their works, the application of which must be done through a tool that they themselves created or modified according to their needs. This is the case of [27], who automates the method for identifying multidimensional concepts in operational sources for the MD. Mazón and Trujillo, in conjunction with Glorio [28], automatically derive standard conceptual models that are constructed with the dimension hierarchies that do not violate summarizability. [29] manage to derive the conceptual and logical model automatically from the data sources and the business requirements. In addition, in different publications we can see the use of transformations among models using standards such as QVT [30].

- (RQ4): The most researched stage is conceptual design with a total of 17 publications. Second we found publications focused on two stages: conceptual and logical design. There are 6 are works that investigate those two areas together. Finally, we found 2 publications oriented towards the stage of logical design. We did not find any studies dealing with guidelines to
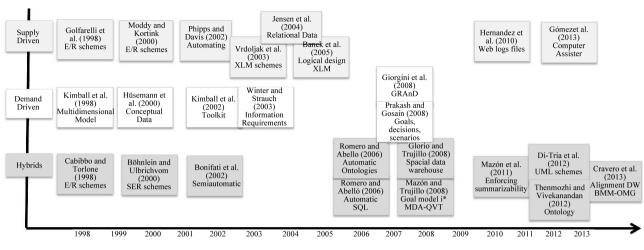
Figure 4. Timeline of the selected works.

create the physical design of the DW.

6. Limitations of the Study

The main threats to the validity of this systematic mapping are related to bias in the selection of the studies that must be included and, in some cases, to the possible inaccuracies in the data extraction, as we are aware that it can happen that some existing documents have not been included, although the broad review that was undertaken and the knowledge of this subject have led us to conclude that if there are any, there are probably not many.

We tried to minimize this problem by performing the following activities before selecting studies: (1) explain what a DW design methodology to persons selected papers, (2) validate that everyone understood the same, through exhibitions, (3) using the kappa index as a validator of the selection process. The second limitation is the quality of the studies incorporated. This limitation could have been avoided by doing a systematic review of these works. However, the low number of truly pertinent publications (25) leads us to believe that it is still soon for this type of quality assessment. Another possible limitation of this type of study is the possibility of being mistaken in the classification due to the ambiguous use that the authors make of it such as, in our case, approach or methodology.

7. Related Work

There several studies comparing different methods and approaches to the design of DW.

List and others authors [10], analyze various DW development methodologies, using criteria such as: end user involvement, duration of development and completion, skill level of data warehouse designer, complexity of data model, amount of source systems and longevity of data model. The target was to establish a link between the methodology and the requirement domain.

On the other hand, Romero and Abelló [5], compared 17 selected methodologies based on 3 criteria: papers with a high number of citations, papers presenting a novel contribution, and in case of papers created by the same authors, have included the latest version. To make the comparison created a common framework describing different features of the DW. They used a methodology for the selection of articles and posterios discucion.

Jindal and Tajena [31], conducted a comparative study small based on the following criteria: Proposal, Framework/Architecture, Approach or technique proposed, schema used, whether the design can be extended to logical and physical design also, case study and tool used. The aim was to propose a generalized object oriented conceptual design framework based on UML that meets all types of user needs.

Finally, Cravero and Sepúlveda [32], perform a chronological study of various methodologies, but there is not a comparative study. They used a systematic methodology for the selection of papers.

All these works present a study on methodologies for DW design, but do not use a methodology recognized by the software engineering community, such as the systematic mapping.

8. Conclusions

In this article, we have presented a systematic mapping of studies of the paradigms for the MD of DW, having provided a framework of current work, helping open the field to new researchers.

The review framework and the protocol used to perform this review guarantee the completeness of the results. In conclusion, the significant deficiency that we identified is the lack of feedback when the methodologies are applied, because developing, starting up and obtaining results in a DW take a minimum of 5 years ap-

proximately, which is why there are no articles that deal with the complete implementation process of a DW either in an academic or industrial environment. Another deficiency that we identified is the lack of experiments in this line of research: there are no publications that address the design and development of a DW by applying and comparing different existing methodologies in the area. The lack of tools is another point that we must emphasize; in most of the publications that contribute with tools, we found only prototypes. There is no instrument finalized and validated by the investigators in the research field and this deficiency is important considering that most of the publications focus on industry, a market that determines the evolution of the area. There is also no finalized tool to realize the conceptual design totally automatically, which is the objective that most of the authors wish to achieve.

Acknowledgements

REFERENCES

[1] P. Giorgini, S. Rizzi and M. Garzetti, "Goal Oriented Requirement Analysis for Data Warehouse Design," *DOLAP*'05, *Proceedings of the 8th ACM International Workshop on Data Warehousing and OLAP*, Bremen, Vol. 45, 2005, pp. 47-56.

[2] R. Kimball and M. Ross, "The Data Warehouse Toolkit," 2nd Edition, John Wiley & Sons, Chichester, 2002.

[3] M. Frolick and K. Winter, "Critical Factors for Data Warehouse Failure," *Business Intelligence Journal*, Vol. 8, No. 1, 2003, pp. 61-71.
http://www.tdwi.org/research/display.aspx?ID=6592

[4] A. Cravero, S. Sepúlveda, J.-N. Mazón and J. Trujillo, "Alineamiento de Objetivos de la Organización: Una Necesidad para el Análisis de Requisitos en Almacenes de Datos," *CIBSE*2010, 2010.

[5] O. Romero and A. Abello, "A Survey of Multidimensional Modeling Methodologies," *International Journal of Data Warehousing & Mining* (*IJDWM*), Vol. 5, No. 2, 2009, pp. 1-23.
http://dx.doi.org/10.4018/jdwm.2009040101

[6] R. Winter and B. Strauch, "A Method for Demand Driven Information Requirements Analysis in Data Warehousing Projects," *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, 2003, pp. 9-14.

[7] B. Kitcheham, T. Dyba and M. Jorgensen, "Evidence-Based Software Engineering," *Proceeding of the 26th International Conference on Software Engineering* (*ICSE*, 2006), IEEE Computer Society, 2006, pp. 373-378.

[8] N. Prakash and A. Gosain, "An Approach to Engineering the Requirements of Data Warehouses," *Requirements Engineering*, Vol. 13, No. 1, 2008, pp. 49-72.
http://dx.doi.org/10.1007/s00766-007-0057-x

[9] S. Rizzi, A. Abelló, J. Lechtenbörger and J. Trujillo, "Research in Data Warehouse Modeling and Design: Dead or Alive?" *Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP*, Arlington, 2006, pp. 3-10.

[10] B. List, R. M. Bruckner, K. Machaczek and J. Schiefer, "A Comparison of Data Warehouse Development Methodologies Case Study of the Process Warehouse," In A. Hameurlain, R. Cicchetti and R. Traunmüller, Eds., *Proceedings of 13th International Conference on Database and Expert Systems Applications*, *Lecture Notes in Computer Science*, Aix-en-Provence, Vol. 2453, 2002, pp. 203-215.

[11] M. Golfarelli, "From User Requirements to Conceptual Design in Data Warehouse Design," *Data Warehousing Design and Advanced Engineering Applications*: *Methods for Complex Construction*, 2010, p. 1.

[12] P. Giorgini, S. Rizzi and M. Garzetti, "GRAnD: A Goal-Oriented Approach to Requirement Analysis in Data Warehouses," *Decision Support Systems*, Vol. 45, No. 1, 2008, pp. 4-21. http://dx.doi.org/10.1016/j.dss.2006.12.001

[13] R. Weir, T. Peng and J. Kerridge, "Best Practice for Implementing a Data Warehouse: A Review for Strategic Alignment," School of Computing, Napier University, Edinburgh, 2003.

[14] W. Inmon, "Building the Data Warehouse," 2nd Edition, 1996.

[15] E. Codd, S. Codd and C. Salley, "Providing OLAP to User-Analysts: An IT Mandate, E. F. Codd and Associates," 1993.

[16] F. Silva-Paim and J. F. B.-Castro, "DWARF: AN Approach for Requirements Definition and Management of Data Warehouse Systems," *Proceedings of the Requirements Engineering Conference*, *RE*'03, 11th IEEE International, 2003, pp. 75-84.

[17] K. Petersen, R. Feldt, S. Mujtaba and M. Mattsson, "Systematic Mapping Studies in Software Engineering," *EASE*'08 *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering. British Computer Society Swinton*, 2008, pp. 68-77.

[18] B. Kitchenham and S. Charters, "Guidelines for Performing Systematic Literature Reviews in Software Engineering," Technical Report EBSE'07, 2007.

[19] K. Gwet, "Inter-Rater Reliability: Dependency on Trait Prevalence and Marginal Homogeneity," *Statistical Methods for Inter-Rater Reliability Assessment Series*, Vol. 2, 2002, pp. 1-9.

[20] P. Hernández, I. Garrigos and J.-N. Mazón, "Model-Driven Development of Multidimensional Models from Web Log Files," *Advances in Conceptual Modeling—Applications and Challenges*, 2010, pp. 170-179.

[21] F. Di-Tria, E. Lefons and F. Tangorra, "Hybrid Methodology for Data Warehouse Conceptual Design by UML Schemas," *Information and Software Technology*, Vol. 54, No.

4, 2012, pp. 360-379.
http://dx.doi.org/10.1016/j.infsof.2011.11.004

[22] J. N. Mazón, J. Pardillo and J. Trujillo, "A Model Driven Goal Oriented Requirement Engineering Approach for Data Warehouses," ER Workshops (RIGiM), 2007.

[23] P. Giorgini, S. Rizzi and M. Garzetti, "Goal Oriented Requirement Analysis for Data Warehouse Design," DO-LAP'05, Bremen, Vol. 45, 2005, pp. 47-56.

[24] A. Bonifati, F. Cattaneo, S. Ceri, A. Fuggetta and F. Paraboschi, "Designing Data Marts for Data Warehouses," *ACM Transactions on Software Engineering and Methodology*, Vol. 10, No. 4, 2001, pp. 452-483.
http://dx.doi.org/10.1145/384189.384190

[25] A. Cravero, J.-N. Mazón and J. Trujillo, "A Business-Oriented Approach to Data Warehouse Development," *Ingeniería e Investigación*, Vol. 33, No. 1, 2013, pp. 59-65.

[26] OMG, "Business Motivation Model. Version 1.1," OMG Document Number: Formal/2010-05-01, 2008.

[27] O. Romero and A. Abelló, "Multidimensional Design Methods for Data Warehousing," *Integrations of Data Warehousing*, *Data Mining and Database Technologies*: *Innovative Approaches*, 2011, p. 78.

[28] O. Glorio, J. Pardillo, J.-N. Mazón and J. Trujillo, "Da-WaRA: An Eclipse Plugin for Usingi* on Data Warehouse Requirement Analysis," RE'08, 2008, pp. 317-318.

[29] M. Thenmozhi and K. Vivekanandan, "An Ontology Based Hybrid Approach to Derive Multidimensional Schema for Data Warehouse," *International Journal of Computer Applications*, Vol. 54, No. 8, 2012, pp. 36-42.

[30] J.-N. Mazón and J. Trujillo, "An MDA Approach for the Development of Data Warehouses," *JISBD*'09, 2009, p. 208.

[31] R. Jindal and S. Taneja, "Comparative Study of Data Warehouse Design Approaches: A Survey," *International Journal of Database Management Systems*, Vol. 4, No. 1, 2012, pp. 33-45.

[32] A. Cravero and S. Sepúlveda, "A Chronological Study of Paradigms for Data Warehouse Design," *Ingeniería e Investigación*, Vol. 32, No. 2, 2012, pp. 58-62.