Journal of Universal Computer Science, vol. 21, no. 11 (2015), 1508-1534
submitted: 10/5/14, accepted: 12/1/15, appeared: 1/11/15 © J.UCS

DWCMM: The Data Warehouse Capability Maturity Model

Marco Spruit
(Institute of Information and Computing Sciences, Utrecht University, The Netherlands
m.r.spruit@uu.nl)

Catalina Sacu
(Institute of Information and Computing Sciences, Utrecht University, The Netherlands
caitlin_shacu@yahoo.com)

Abstract: Data Warehouses and Business Intelligence have become popular fields of research in recent years. Unfortunately, in daily practice many Data Warehouse and Business Intelligence solutions still fail to help organizations make better decisions and increase their profitability, due to intransparent complexities and project interdependencies. In addition, emerging application domains such as Mobile Learning & Analytics heavily depend on a well-structured data foundation with a longitudinally prepared architecture. Therefore, this research presents the Data Warehouse Capability Maturity Model (DWCMM) which encompasses both technical and organizational aspects involved in developing a Data Warehouse environment. The DWCMM can be used to help organizations assess their current Data Warehouse solution and provide them with guidelines for future improvements. The DWCMM consists of a maturity matrix and a maturity assessment questionnaire with 60 questions. The DWCMM has been evaluated empirically through expert interviews and case studies. We conclude that the DWCMM can be successfully applied in practice and that organizations can intelligibly utilize the DWCMM as a quickscan instrument to jumpstart their Data Warehouse and Business Intelligence improvement processes.

Keywords: Data Warehousing, Business Intelligence, Maturity Modelling, Mobile Analytics
Categories: D.2.8, H.1.1

1 Introduction and Problem Definition

In nowadays economy, organizations are part of a very dynamic environment due to continuous changing conditions and relationships. As Kaye (1996, p. 20) notes, "organizations must collect, process, use, and communicate information, both external and internal, in order to plan, operate and take decisions". The ongoing request for profits, increasing competition and demanding customers, all require organizations to take the best decisions as fast as possible (Vitt & Luckevich, 2002). One of the solutions that can narrow down the period of time between the moment of acquiring the information and getting the right results to improve the decision making process is the implementation of Data Warehouses and Business Intelligence (BI) applications.

Over the years, data warehouses (DWs) and BI solutions have become one of the fundamentals of the information systems that are used to support the decision making initiatives. Most large companies have already established DW systems as a component of the information systems landscape. According to Gartner (2007) BI and

DWs are at the forefront of the use of IT to support management decision-making. DWs can be thought of as the large-scale data infrastructure for decision support. BI can be viewed as the data analysis and presentation layer that sits between the DW and the executive decision-makers (Arnott & Pervan, 2005). In this way, the DW/BI solutions can transform raw data into information and then into knowledge.

However, a DW is not only a software package. The adoption of DW technology requires massive capital expenditure and a certain deal of implementation time. DW projects are hence very expensive, time-consuming and risky undertakings compared with other information technology initiatives, as cited by prior researchers (Wixom & Watson, 2001; Gartner, 2007; Solomon, 2005). Moreover, it is often believed that one-half to two-thirds of all initial DW efforts fail (Hayen, et al., 2007). Gartner (2007) estimates that more than fifty percent of DW projects have limited acceptance or fail. Therefore, it is crucial to have a thorough understanding of the critical success factors and variables that determine the efficient implementation of a DW solution.

These factors can refer to the development of the DW/BI solution or to the usage and adoption of BI. In this research, we will focus on the former as we consider that it represents the foundation for a solid DW solution that can have a high rate of usage and adoption. First, it is critical to properly design and implement the databases that lie at the heart of the DW. The right architecture and design can ensure performance today and scalability tomorrow. Second, all components of the DW solution (*e.g.* data repository, infrastructure, user interface) must be designed to work together in a flexible, easy-to-use way. A third task is to develop a consistent data model and establish what and how source data will be extracted. In addition to these factors, the DW needs to be created and developed quickly and efficiently so that the organization can gain the business benefits as soon as possible (AbuAli & Abu-Addose, 2010). As can be seen, a DW project can unquestionably be complex and challenging, and there is usually not a single successful solution that can be applied to all organizations. Therefore, it is very important for organizations to be aware of their current situation and know the steps they need to take for continuous improvement. However, an objective assessment often proves to be a difficult task.

Maturity models can be helpful in this situation. They essentially describe the development of an entity over time, where the entity can be anything of interest: a human being, an organizational function, an organization, and so on. (Klimko, 2001). Maturity models have a number of sequentially ordered levels, where the bottom stage stands for an initial state than can be, for example, characterized by an organization having little capabilities in the domain under consideration. In contrast, the highest stage represents a conception of total maturity. Advancing on the evolution path between the two extremes involves a continuous progression regarding the organization's capabilities or process performance. The maturity model serves as an assessment of the position on the evolution path, as it offers a set of criteria and characteristics that need to be fulfilled in order to reach a particular maturity level (Becker, et al., 2009).

With the help of maturity modelling, we will gain some insight into the technical and organizational variables that determine the successful development of a DW solution and analyze these variables. Therefore, in order to make an assessment of the most important aspects that influence a DW project, this paper develops a *Data Warehouse Capability Maturity Model* (DWCMM**)** which provides an answer to the

following research question: *How can the maturity of a company's data warehouse technical aspects be assessed and acted upon?*

2 Research Methodology

The main goal of this research is to develop a DWCMM that depicts the maturity stages of a DW project. For this purpose, a design research approach is used as its main philosophy is to generate scientific knowledge by building and validating a previously designed artifact (Hevner, et al., 2004). In this research, the artifact is the DWCMM, which is developed according to the five steps in developing design research artifacts as described in (Vaishnavi & Kuechler, 2008): problem awareness, suggestion and development, evaluation and conclusion. Awareness of the problem was raised in discussions with DW/BI practitioners and literature study on data warehousing and maturity modelling. A detailed problem description was provided in the section before. Based on this, it has become clear that DW projects often fail or do not bring the expected results and that organizations sometimes need guidelines for improvement.

As a solution to this problem, we developed the DWCMM which can be used to assist organizations in doing a maturity assessment for the DW technical aspects and in providing guidelines for future improvements. First, an overview on the model and its main components will be presented in section 3. We will then elaborate on each category of the DWCMM and each part of the maturity assessment questionnaire in sections 4 and 5. The results of the evaluation phase are presented in section 6. The DWCMM has been evaluated by carrying out five expert interviews and multiple case studies within four organizations, following Yin's (2009) case study approach. Finally, section 7 provides conclusions regarding our model and agenda for future research.

3 Towards The Data Warehouse Capability Maturity Model

In literature, a lot of maturity models have been developed (de Bruin, Freezey, Kulkarniz, & Rosemann, 2005), but only some of them managed to gain global acceptance. There are also several information technology and/or information system maturity models dealing with different aspects of maturity: technological, organizational and process maturity. Some of them are specific to the data warehousing/BI field. The most important maturity models that served as a source of inspiration for our research can be seen in table 1.

Each of these models has a different way of assessing maturity, but there are some common components for all the models. All the models have interesting elements, but also weak points that could be improved. Moreover, all the models developed for the field of data warehousing/BI focus on more variables involved in such a project, but they do not go deep into analyzing the technical aspects.

Authors	**Model**	**Focus**
Nolan (1973)	Stages of Growth	IT Growth Inside an Organization
Software Engineering Institute (1993)	Capability Maturity Model (CMM)	Software Development Processes
Watson, Ariyachandra & Matyska (2001)	Data Warehousing Stages of Growth	Data Warehousing
Chamoni & Gluchowski (2004)	Business Intelligence Maturity Model	Business Intelligence
The Data Warehousing Institute (2004)	Business Intelligence Maturity Model	Business Intelligence
Gartner – Hostmann (2007)	Business Intelligence and Performance Management Maturity Model	Business Intelligence, Performance Management

Table 1: An overview of maturity models.

The maturity model which served as the main foundation for this research is the CMM (Paulk, Weber, Curtis, & Chrissis, 1995). It has become a recognized standard for rating software development organizations. The CMM is a framework that describes the key elements of an effective software process and presents an evolutionary improvement path from an ad-hoc, immature process to a mature, disciplined one. Since its development, CMM has become a universal model for assessing software process maturity. However, the CMM has often been criticized for its complexity and difficulty of implementation. That is why we simplified it by keeping the five maturity levels—*i.e.* initial, repeatable, defined, managed and optimizing—the process capabilities and the key process areas, which in our model would translate to the chosen benchmark variables/categories for doing the DW maturity assessment.

Therefore, it can be seen that even if DW/BI solutions are often implemented in practice and a lot of maturity models have been created, none is actually focusing on the technical aspects of the DW/BI solution and the organizational processes that sustain them. Hence, this is the research gap we would like to fill in by developing a ***Data Warehouse Capability Maturity Model (DWCMM)*** that focuses on the DW technical solution and DW organization and processes. The DWCMM can be depicted in figure 1. A short overview of the model and its components will be provided in the next paragraphs.

When analyzing the maturity of a DW solution, we are actually taking a snapshot of an organization at the current moment in time. Therefore, in order to do a valuable assessment, it is important to include in the maturity analysis the most representative dimensions involved in the development of a DW solution. Several authors describe that the main phases usually involved in a DW project lifecycle are: project planning and management, requirements definition, design, development, testing and acceptance, deployment, growth and maintenance (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008; Moss & Atre, 2003; Ponniah, 2001). All of these phases and

processes refer to the implementation and maintenance of the actual DW technical solution which includes: the general architecture and infrastructure, data modelling, ETL, BI applications. These categories can be analyzed from many points of view which will be depicted in our model and the maturity assessment we developed. Therefore, the DWCMM will be restricted for doing the assessment of the technical aspects, without taking into consideration the DW/BI usage and adoption or the DW/BI business value. It will consider two main benchmark variables/categories for analysis, each of them having several sub-categories. Firstly, the **DW Technical Solution** consists of the following four components: General Architecture and Infrastructure, Data Modelling, Extract-Transform-Load (ETL) and BI Applications. Secondly, the **DW Organization & Processes** dimension comprises the following two aspects: Development Processes and Service Processes.

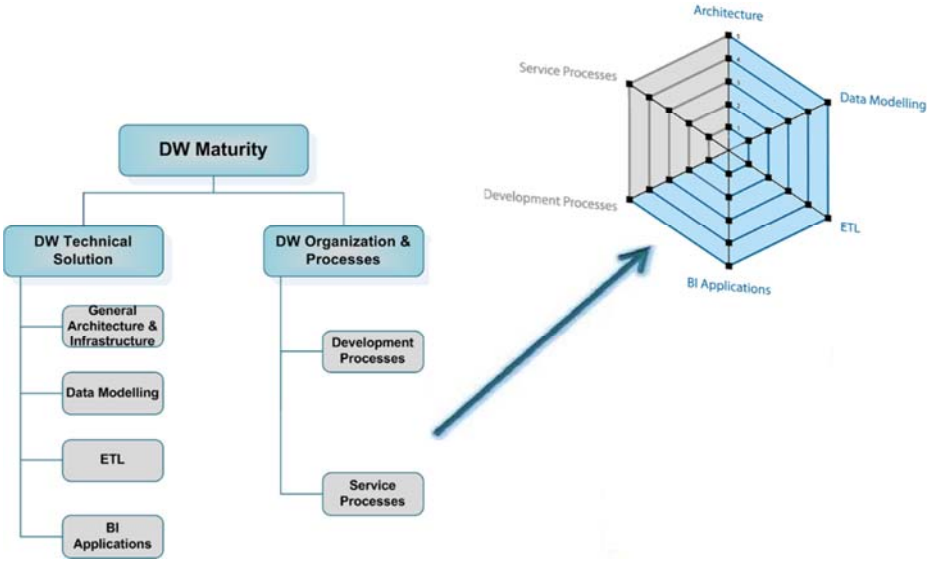

Figure 1: The Data Warehouse Capability Maturity Model (DWCMM) categories.

As can be seen from figure 1, the DWCMM does a maturity assessment which will provide a maturity score for each benchmark sub-category. In order to create a complete image on the current DW solution for an organization, the DWCMM has several components that will be described further in this section: a maturity assessment questionnaire, a condensed maturity matrix, and a complete maturity matrix.

3.1 Maturity assessment questionnaire

The complete DW maturity assessment questionnaire has been published as an Utrecht University technical report (Sacu, Spruit, & Habers, 2010). Emphasis should be put on two aspects regarding the DW maturity assessment questionnaire. Firstly, it does *a high level assessment of an organization's DW solution* and it is *limited strictly to the DW technical aspects*. Secondly, the model will assess "*what*" and "*if*" certain

characteristics and processes are implemented and not *"how"* they are implemented. The DW maturity assessment questionnaire has 60 questions divided into the following three categories: DW General Questions, DW Technical Solution, and DW Organization & Processes.

DW General Questions (9 questions) – it comprises of several questions about the DW/BI solution and they are not scored. Their purpose is to offer a better image on the drivers for implementing the DW environment, the budget allocated for data warehousing and BI, the DW business value, end-user adoption, etc. This will be useful in creating a complete picture on the current DW solution and its maturity. Also, once the questionnaire is filled in by more organizations, this data will serve as input for statistical analysis and comparisons between organizations from the same industry or across industries.

DW Technical Solution (32 questions) – it comprises of several scored questions for each of the following sub-categories: General Architecture and Infrastructure (9 questions), Data Modelling (9 questions), ETL (7 questions), and BI Applications (7 questions). More details on this part will be given in the next sections.

DW Organization & Processes (19 questions) – it comprises of several scored questions for each of the following sub-categories: Development Processes (11 questions), and Service Processes (8 questions). Again, more details on this part will be given in the next sections.

Each question from the questionnaire will have five possible answers which are scored from 1 to 5, 1 being a characteristic for the lowest maturity stage and 5 for the highest one. When an organization takes the survey, it will first receive a *maturity score for each sub-category* by computing the average value of the weightings (*i.e.*: sum of the weightings / number of questions); then, *an overall score for each of the two main categories* will be given by computing the average value of the scores obtained for each sub-category; and finally, *an overall maturity score* is shown following the same principle applied to the main two categories scores.

We believe that the *maturity scores for the sub-categories* can give a good overview on the current DW solution implemented by the organization. This is the reason why, after computing the maturity scores for each sub-category, a radar graph as the one depicted in figure 1 will be drawn to show the alignment between these scores. In this way, the organization will have a clearer image on their current DW project and will know what sub-category is the strongest and which one is left behind.

Moreover, after reviewing the maturity scores and the given answers by a specific organization, some *general feedback and advice for future improvements* will be provided. Each organization that takes the assessment will receive a document with a short explanation on the scoring method, a table with their maturity scores and the radar graph, and then some general feedback that will consist of: a general overview on the maturity scores; an analysis of the positive aspects already implemented in the DW solution; and several steps that the organization should take in order to improve their current DW application.

3.2 Condensed DW maturity matrix

As our model measures the maturity of a DW solution, we also created two maturity matrices – a condensed maturity matrix and a detailed one – each of them having five maturity stages as inspired by the CMM: Initial (1); Repeatable (2); Defined (3);

Managed (4); Optimized (5); where the initial stage describes an incipient DW development and the optimized level shows a very mature solution that can be obtained by an organization with a lot of experience in the field where everything is standardized and monitored. An organization will usually be situated on different stages of maturity for each sub-category that will determine the overall maturity level.

The condensed DW maturity matrix gives a short overview of the most important characteristics for each sub-category for each maturity level. This will offer a better image on the main goal of the DWCMM and on what the detailed maturity matrix entails. The condensed maturity matrix can be seen in Table 2.

	Stages / *Categories*	Initial (1)	Repeatable (2)	Defined (3)	Managed (4)	Optimized (5)
DW TECHNICAL SOLUTION	**Architecture**	Desktop data marts	Independent data marts	Independent data warehouses	Central DW with/ without data marts	DW/BI service that federates a central DW and other sources via standard interface
	Data Modelling	No data models synchroni-zation or standards	Manually synchronized data models	Manually or automatically synchronized data models	Automatic synchronization of most data models	Enterprise-wide standards and automatic synchronization of all the data models
	ETL	Simple ETL with no standards that just extracts and loads data into the DW	Basic ETL with simple transformations	Advanced ETL (*e.g.* slowly changing dimensions manager, data quality system, reusability, etc.)	More advanced ETL (*e.g.* hierarchy manager, special dimensions manager, etc.)	Optimized ETL for real-time DW with all the standards defined
	BI Applications	Static and parameter-driven reports	Ad-hoc reporting; OLAP	Dashboards & scorecards	Predictive analytics; data & text mining	Closed-loop & real-time BI applications
DW ORGANIZATION & PROCESSES	**Development Processes**	Ad-hoc, non-standardized development processes or defined phases	Some development processes policies and procedures established with some phases separated	Standardized development processes with all the phases separated and all the roles formalized	Quantitative development processes management	Continuous development processes improvement
	Service Processes	Ad-hoc, non-standardized service processes	Some service processes policies and procedures established	Standardized service processes with all the roles formalized	Quantitative service processes management	Continuous service processes improvement

Table 2: The DWCMM Condensed Maturity Matrix.

3.3 Complete DW maturity matrix

We will give a short overview on the detailed DW maturity matrix in this paragraph. The complete DWCMM is included in the Appendix A. First, the characteristics for each maturity stage are usually obtained by mapping the correspondent answers of each question from the maturity assessment questionnaire (except for several characteristics such as: project management, testing and acceptance, whose answers are formulated in a different way). In this way, an organization will be able to see their maturity stage by category (*e.g.* General Architecture and Infrastructure) and by main category characteristics (*e.g.* metadata, standards, infrastructure, etc.). The matrix has two dimensions: columns and rows.

Columns show each benchmark sub-category (*i.e.*: General Architecture and Infrastructure, Data Modelling, ETL, BI Applications; Development Processes, Service Processes) with their maturity stages from Initial (1) to Optimized (5). *Rows* show the main analyzed characteristics (*e.g.* for General Architecture and Infrastructure – conceptual architecture, business rules, metadata, security, data sources, performance, infrastructure, update frequency) for each sub-category divided by maturity stage.

Moreover, the matrix can be interpreted in two ways. First, one could take each stage and see the specific characteristics for each sub-category for that particular stage. Second, one could take each sub-category and see the specific characteristics for each stage or for a particular stage.

As the developed questionnaire does an assessment for each benchmark sub-category, a specific organization will most likely follow the second interpretation. They would probably like to know what steps to take to improve each sub-category and hence, the overall maturity score, which will lead to a higher maturity stage. It is also very unlikely that an organization will have all the characteristics for all the sub-categories on the same maturity stage at the same moment in time. Therefore, if a company gets a maturity score of 3, this does not mean that all the characteristics for all the sub-categories are on stage three. Depending also on the standard deviation and the answers themselves, we can find out more information about the actual situation.

Now that the main components of the DWCMM have been identified, we will continue by taking a closer look at the main categories and sub-categories of the model and their analyzed characteristics. These can be depicted in the maturity assessment questionnaire and detailed maturity matrix. We will start with the DW technical solution and continue with the DW organization and processes.

4 DW Technical Solution Maturity

As mentioned earlier, the main components that need to be analyzed when doing an assessment of the DW technical solution are: general architecture and infrastructure, data modelling, ETL and BI applications.

4.1 General Architecture and Infrastructure

DW *architecture* includes: three main components (*i.e.*: data modelling, ETL, BI applications), several data storage components (*e.g.* source systems, data staging area,

DW database, operational data store, data marts) and the way they are assembled together (Ponniah, 2001), and underlying elements such as infrastructure, metadata and security that support the flow of data from the source systems to the end-users (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008; Chauduri & Dayal, 1997). This is connected to the conceptual approach of designing and building the DW (*e.g.* conformed data marts – Kimball or enterprise-wide DW – Inmon, etc.). Therefore, in this research we consider architecture and infrastructure as a separate sub-category for assessing maturity and for which the main characteristics will be further analyzed.

Conceptual architecture and its layers (question 1) – encompasses the conceptual approach of designing and building the DW with all its data storage layers.

DW data sources (question 6) - the types of data sources that the DW extracts data from (*e.g.* Excel files, text files, relational databases, ERP & CRM systems, unstructured data: text documents, e-mails, images, videos, Web data sources).

Infrastructure (question 8) – it provides the underlying foundation that enables the DW architecture to be implemented (Ponniah, 2001), and it includes elements such as: hardware platforms and components, operating systems, database platforms, connectivity and networking (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008).

Metadata management (question 4) – metadata can be seen as all the information that defines and describes the structures, operations and contents of the DW system in order to support the administration and effective exploitation of the DW. The main elements that influence its maturity are: the types of implemented metadata (*i.e.*: business, technical or process) and the integration of metadata repositories (Moss & Atre, 2003; Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008).

Security management (question 5) – user access security is usually implemented through several methods, presented here in hierarchical order of difficulty of implementation (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008; Moss & Atre, 2003; Ponniah, 2001): authentication, tool-based security, role-based security, authorization.

Business rules (questions 2 & 3) – they are abstractions of the policies and practices of a business organization (Kaula, 2009), and are used to capture and implement precise business logic in processes, procedures, and systems (manual or automated).

Performance optimization (question 7) – encompasses the various methods needed to improve DW performance (Ponniah, 2001): software performance improvement (*e.g.* index management, data partitioning, parallel processing, view materialization); hardware performance improvement; specialized DW appliances or cloud computing which are characteristics for a very high stage of maturity.

Update frequency (question 9) – it is one of the characteristics that differentiate classical DW solutions built for strategic and tactical BI from the newer DWs that process data in real time.

4.2 Data Modelling

Data modelling is the process of creating a data model. *A data model* is "a set of concepts that can be used to describe the structure of and operations on a database" (Navathe, 1992). Data modelling is very important for creating a successful information system as it defines not only data elements, but also their structures and

relationships between them. The most important characteristics which should be taken into consideration when assessing the maturity of data modelling are described below.

Synchronization between all the data models found in the DW (question 2) – establishing consistency among data from a source to a target data storage and vice versa and the continuous harmonization of the data over time.

Design levels (question 3) – encompasses all the data model design levels: conceptual design, logical design and physical design.

Tool (question 1) – data models can be created by just drawing the models in different spreadsheets and documents. However, the more mature solution is to use a data modelling tool that can make the design itself and metadata management easier and more efficient.

Standards (questions 4 & 5) – standards in a DW environment are necessary and cover a wide range of objects, processes, and procedures. All the maturity assessments related to standards will address general aspects such as the definition and documentation of standards and their actual implementation. Most often, standards related to data modelling refer to naming conventions for the objects and attributes in the data models.

Metadata management (question 6) – encompasses the common subset of business and technical metadata components as they apply to data (Moss & Atre, 2003): data names, definitions, relationships, identifiers, types, lengths, policies, ownership, etc.

Dimensional modelling (questions 7, 8 & 9) – there are several data modelling techniques that can be applied for data warehousing: relational (or normalized), dimensional, data vault, etc. In this research we focused on dimensional modelling. See (Kimball, 1996) for more information on dimensional modelling.

4.3 Extract-Transform-Load (ETL)

As the name shows, the *Extract-Transform-Load (ETL)* process mainly involves the following activities: extracting data from outside sources; transforming data to fit the target's requirements; loading data into the target database. The ETL system is very complex and resource demanding (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008), and hence, 60 to 80 percent of the time and effort of developing a DW project is devoted to the ETL system (Nagabhushana, 2006). The main characteristics that we included in our ETL maturity assessment are further described in this paragraph.

Complexity (question 2) – this refers to the maturity and performance of each ETL component (*i.e.*: extract, transform, load). For example, the extraction phase should include a data profiling system, a change data capture system and the extract system itself. The transformation step usually includes cleaning and transforming data according to the business rules and standards that have been established for the DW. The DW load system takes the load images created by the extraction and transformation subsystems and loads these images directly into the DW.

Data quality system (question 3) – data quality is critical for the success of a DW. Therefore, we decided to include a question that would depict its main characteristics for each maturity stage regarding: daily automation, specific data quality tools, identifying data quality issues and actually solving them.

Management and monitoring (question 4) – encompasses all the necessary capabilities for the ETL processes to run consistently to completion and be available

when needed (*e.g.* an ETL job scheduler; a backup system; a recovery and restart system – it can be manual or automatic; a workflow monitor, etc.)

Tool (question 1) – there is a constant debate whether an organization should deploy custom-coded ETL solutions or should buy an ETL tool suite (Kimball & Caserta, 2004). A company that uses hand-coded ETL usually does not have a very complex ETL process which shows a low level of maturity regarding ETL capabilities.

Metadata management (question 7) – ETL is responsible for the creation and use of much of the metadata describing the DW environment. Therefore, it is important to capture and manage all possible types of metadata for ETL: business, technical and process metadata.

Standards (questions 5 & 6) – includes ETL specific standards that are related to: naming conventions, set-up standards, recovery and restart system, etc.

4.4 BI Applications

BI applications, sometimes referred to as "front-end" tools (Chauduri & Dayal, 1997), are what the end-users see and hence, are very important for a DW to be considered a successful one. According to March & Hevner (2007), a crucial point for achieving DW implementation success is the selection and implementation of appropriate end-user analysis tools, because business benefits of BI are only gained when the system is adopted by its intended end-users. The main aspects that determine the maturity of BI applications are analyzed further in this paragraph.

Types of BI applications (question 1) – encompasses the main types of BI whose complexity contributes to the maturity of a DW environment. According to Azvine (2005), traditional BI applications fall into the following categories sorted by ascending complexity: *report what has happened* – standard reporting and query applications; *analyze and understand why it has happened* – ad-hoc reporting and online analytical processing (OLAP); visualization applications (*i.e.*: dashboards, scorecards); *predict what will happen* – predictive analytics (*i.e.*: data and text mining). In the last couple of years, due to the development of real-time data warehousing, a new category of BI applications – operational BI and closed-loop applications – has developed (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008).

Delivery method (question 6) – it includes the main BI applications delivery methods. As end users are interested only in the results they get from the BI applications, the easiness of accessing and delivering these results is critical for the success of the DW solution.

Tool (question 2) – defines the usage of BI applications tools which can really make a difference for the DW solution.

Metadata management (question 7) – encompasses the main metadata accessibility methods. As BI applications are what the end user sees, this is an important aspect for DW success (Moss & Atre, 2003).

Standards (questions 3 & 4) – it includes standards specific to BI Applications such as: naming conventions, generic transformations, logical structure of attributes and measures, etc.

5 DW Organization and Processes Maturity

When assessing the maturity of a DW technical solution, the processes and roles involved in the project also need to be analyzed. A good technical solution cannot be developed without the processes surrounding it as there is a strong interconnection between the two parts. The necessary processes for a DW project are: development processes and service processes.

5.1 DW Development Processes

A DW solution can be considered a software engineering project with some specific characteristics. And, therefore, as any software engineering project, it will go through several development stages (Moss & Atre, 2003). Since DW/BI is an enterprise-wide evolving environment that is continually improved and enhanced based on feedback from the business community, the best approach for its development is *iterative and incremental development*, with *agile techniques* for the development of BI applications (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008; Ponniah, 2001). The high level phases and tasks required for an effective DW implementation are (Kimball, Ross, Thornthwaite, Mundy, & Becker, 2008; Moss & Atre, 2003): project planning and management; requirements definition; design; development; testing and acceptance; deployment/production. The main characteristics which might influence the maturity of DW development processes can be seen below.

CMM levels (question 1) – as it is hard to judge which software development paradigm is better and more mature, the first maturity question on development processes is a more general one and it refers to how the DW development processes map to the CMM levels.

Project planning and management (question 7) – encompasses the main elements that determine the maturity of this characteristic: project planning and scheduling; project risk management; project tracking and control; standard procedure and documentation; and evaluation and assessment (Lewis, 2001).

DW/BI sponsor (question 6) – defines the extent of organizational support and sponsorship for the DW environment. Strong support and sponsorship from senior business management is critical for a successful DW initiative (Ponniah, 2001).

DW project team and roles (question 8) – encompasses how DW project roles and responsibilities are formalized and implemented to solve skill-role mismatches (Humphries, Hawkins, & Dy, 1999).

Requirements definition (question 10) – encompasses how requirements definition is done. In a DW, users' business requirements represent the most powerful driving force (Ponniah, 2001) as they impact virtually every aspect of the project.

Testing and acceptance (question 11) – this is a critical phase for DW success as it includes several important activities which are not always implemented. The degree of implementation influences the success of a DW project and hence, its maturity.

Development/ testing/ acceptance/ production environments (question 2) – encompasses the way organizations set up different environments for different purposes to support all the development phases (Moss & Atre, 2003).

DW quality management (question 5) – its purpose is to provide management with appropriate visibility into the development process being used by the DW project and the products being built (Paulk, Weber, Curtis, & Chrissis, 1995).

Knowledge management (question 9) – encompasses all the knowledge management activities and the way they are implemented.

Standards (questions 3 & 4) – makes an analysis of the standards used for successfully developing, testing and deploying DW functionalities.

5.2 DW Service Processes

In the last two decades, software maintenance began to be treated as a sequence of activities and not as the final stage of a software development project (April, Hayes, Abran, & Dumke, 2004). These processes are very important after a DW has been deployed in order to keep the system up and running and to manage all the necessary changes. Lately, IT organizations made a transition from being pure technology providers to being service providers. This service oriented perspective on IT organizations can be best applied to the software maintenance field as it is an ongoing activity as opposed to the software development which is more project based (Niessink & van Vliet, 2000). Over the years, various IT service frameworks have been proposed, but one that acts as the de-facto standard for the definition of best practices and processes for service support and service delivery is the Information Technology Infrastructure Library (ITIL) (Salle, 2004). Therefore, we will consider the service components from ITIL as a starting point for our analysis of the DW service processes part. Moreover, two maturity models related to IT maintenance and service also served as a foundation for this part of our DW maturity model: the Software Maintenance Maturity Model (April, Hayes, Abran, & Dumke, 2004) and the IT Service CMM (Niessink & van Vliet, 1999). Taking into consideration these models and the changing nature of a DW, we considered the following components when assessing the maturity of DW service processes.

Service quality management (question 2) – this is similar to the DW quality management, but applied to the service processes.

Knowledge management (question 3) – this is also similar to the knowledge management for the DW development processes, but in the context of service processes.

Service level management (question 4) – it negotiates service level agreements (SLAs) with the suppliers and customers and ensures that they are met by continual monitoring and reviewing (Cater-Steel, 2006).

Incident management (question 5) – its main objective is to provide continuity by restoring the service in the quickest way possible by whatever means necessary (Salle, 2004).

Change management (question 6) – it is described as a regular task for immediate and efficient handling of changes that might occur in a DW environment.

Technical resource management (question 7) – the purpose of resource management is to maintain control of the necessary hardware and software resources needed to deliver the agreed DW services level targets (Niessink & van Vliet, 1999).

Availability management (question 8) – manages risks and ensures that all DW infrastructure, processes, tools and roles are according to the SLAs by using appropriate means and techniques (Colin, 2004).

Release management (question 9) – as a DW is continuously changing and evolving over time, the objective of release management is to ensure that only authorized and correct versions of DW are made available for operation (Salle, 2004).

6 Evaluation of the DWCMM

In order to validate the DWCMM, two methods were chosen – expert validation and multiple case studies – on which we will elaborate in this section.

6.1 Expert Validation

To evaluate the utility and further revise the DWCMM, expert validation was applied. An "expert" is defined by Hoffman et al. (1995, p. 132) as a person "highly regarded by peers, whose judgements are uncommonly accurate and reliable and who can deal effectively with rare or tough cases. Also, an expert is one who has special skills or knowledge derived from extensive experience with subdomains". Therefore, eliciting knowledge from experts is very important and useful and can be done using several methods, including structured or unstructured interviews (Hoffman, Shadbolt, Burton, & Klein, 1995).

Experts					
ID	**1**	**2**	**3**	**4**	**5**
Job Position	CI/BI consultant	Principal consultant/ Thought leader BI/CRM	BI consultant	Principal consultant BI	BI consultant
Affiliations					
Industry	DW/BI Consulting	IT Services	BI Consulting	IT Services	DW Consulting
Market	B2B	B2B	B2B	B2B	B2B
Employees	≈ 45	≈ 49000	≈ 35	≈ 38000	≈ 1

Table 3: Experts overview.

Moreover, five experts in data warehousing and BI were interviewed and asked to give their opinions about the content of the model we have developed. The interviews were structured, but consisted of open questions, in order to capture the knowledge of respondents. This offered the possibility of enabling the experts to liberally state their opinions and ideas for improvement. The expert panel consists of five experts from practice, each of them having at least 10 years of experience in the DW/BI field. An overview of the experts and their affiliations is depicted in table 3. All of them are DW/BI consultants at different organizations in The Netherlands (local or multinational).

The experts were asked to give their opinions regarding the DWCMM structure, the DWCMM condensed maturity matrix and the DW maturity assessment questionnaire. All reviewers gave positive feedback for their first impression of all

three deliverables, said they made sense and the model could be applied for assessing an organization's current DW solution. Valuable insights and criticism were provided that resulted in several (mostly minor) improvements. Furthermore, the category "Architecture" was renamed "General Architecture and Infrastructure" as the former created some confusion among the interviewees. Some adjustments were made to the ETL characterization for each stage of the DWCMM condensed maturity matrix. However, most feedback was received regarding the maturity assessment questionnaire. This resulted in two categories of changes: proposed changes that due to time constraints and scope limitation were not implemented in the final version of the model, but should be considered for future research; and implemented improvement suggestions that involved some question rephrasing and answer rephrasing or changing.

6.2 Multiple Case Studies

Depending on the nature of a research topic and the goal of a researcher, different research methods (qualitative and quantitative) are appropriate to be used (Benbasat, Goldstein, & Mead, 1987; Yin, 2009). One of the most widely used qualitative research methods in information systems (IS) research is case study research. It can be used to achieve various research aims: provide descriptions of phenomena, develop theory and test theory (Darke, Shanks, & Broadbent, 1998). In our research, we will use it to test theory which in this case is the DWCMM we developed. The theory is usually either validated or found to be inadequate in some way, and may then be further refined on the basis of the case study findings. Case study research may adopt single or multiple case designs.

As according to Benbasat et al. (1987) and Yin (2009), multiple case studies are preferred over single ones to get better results and analytic conclusions, we decided to conduct a multiple case study research following Yin's (2009) case study approach. In this way, we can achieve a multiple goal: test the model in practice to see if the chosen benchmark variables/categories, the maturity assessment questions and answers match the organizations' specific solutions; and receive feedback and knowledge from respondents regarding the DWCMM in order to make future improvements. Despite the fact that all individual cases are interesting, this section focuses on the overall results.

Organization	**A**	**B**	**C**	**D**
Industry	Retail	Insurance	Retail	Maintenance & Servicing
Market	B2C	B2B & B2C	B2C	B2B
Revenue	19.94 billion €	4.87 billion €	780 million €	NA
Employees	≈ 138000	≈ 4500	≈ 3660	≈ 3500
Respondent Function	BI consultant	DW/BI technical architect	BI manager	BI consultant & DW lead architect

Table 4: Case and Respondent Overview.

Benchmark Category	Organization A	Organization B	Organization C	Organization D
Architecture	2.67	2.56	3.89	3.55
Data Modelling	2.17	3.44	3.00	4.11
ETL	3.14	3.29	3.71	2.86
BI Applications	2.71	2.71	3.43	3.57
Development Processes	2.90	3.19	3.66	3.02
Service Processes	2.63	3.00	2.87	3.12

Table 5: Organizations' Maturity Scores per benchmark category.

Case Overview – The case studies have been conducted at four organizations of different sizes, operating in several types of industries and offering a wide variety of products and services. An overview of the case study organizations (figures are taken from 2009 annual reports) and respondents is depicted in table 4. The main criterion used in the search for suitable organizations was that all approached organizations had a professionally DW/BI system in place whose maturity could be assessed by applying the DWCMM. Furthermore, an important criterion for the selection of respondent per case was that the interviewed respondents had an overall view on the technical and organizational aspects for the DW/BI solution implemented in their organization. A short analysis on the maturity scores each organization got after taking the assessment is also given further in this paragraph.

Case Study Analysis – In this section, a short analysis of the results gotten by all the organizations after filling in the assessment questionnaire is given. The maturity scores regarding the implemented DW solution obtained by the organizations can be seen in the table below.

As shown in the picture depicting our model, a better way to see the alignment between the maturity scores for the six categories is by drawing the radar graph. We will show here the radar graph for organization A as an example.

Some more information regarding the maturity scores for all the four case studies are provided in table 6. As can be seen from table 5, maturity scores for each sub-category are usually between 2 and 4, with one exception: organization D scored 4.11 for Data Modelling. Thus, the overall maturity scores and the total scores per category ranged between 2 and 4 which shows that most organizations are probably somewhere between the second and fourth stage of maturity. The highest maturity score was gotten by organization C, and the lowest one by organization A. Apparently, an overall score close to 4 or 5 is quite difficult to achieve. This is usually normal in maturity assessments, as in practice, nobody is so close to the ideal situation. It will be interesting to see the range of scores after the questionnaire will be filled in by a large number of organizations.

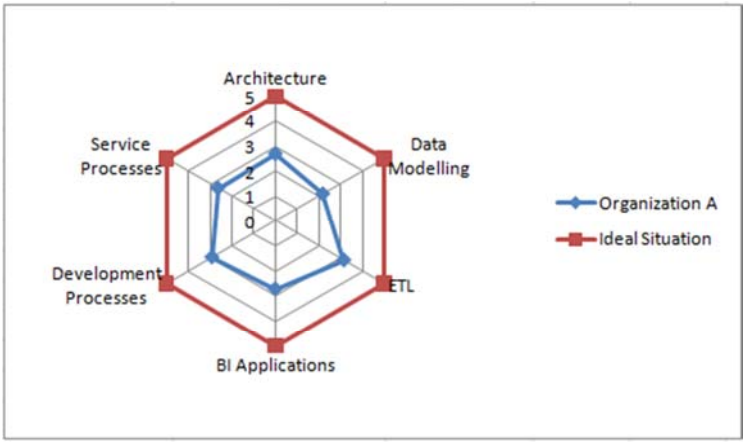

Figure 2: Alignment Between Organization A's Maturity Scores.

Organization Maturity Score	**A**	**B**	**C**	**D**
Total Score for Technical Solution	2.67	3.00	3.51	3.52
Total Score for Org. & Processes	2.77	3.10	3.26	3.07
Overall Score	2.72	3.05	3.38	3.29
Highest Score	3.14	3.44	3.89	4.11
Best DW Category	ETL	Data Modelling	Architecture	Data Modelling
Lowest Score	2.17	2.56	2.87	2.86
Worst DW Category	Data Modelling	Architecture	Service Processes	ETL

Table 6: Maturity Scores Analysis.

From table 6 it can be seen that the categories with the highest and lowest scores are diverse depending on the organization. For example, organization A scored lowest for Data Modelling, whereas Data Modelling was the most mature variable for organization D. Interesting conclusions can also be drawn if comparing the scores for organizations A and C as they are part of the same industry. The former is an international food retailer and has more experience in this industry, whereas the latter is a local one with less experience. However, organization A got a quite low DW maturity score. Thus, experience in the industry does not also mean maturity in data warehousing. Of course, more factors can influence this difference in scores: size, the way data warehousing/BI is embedded in the organizational culture, the percentage from the IT budget for BI, etc.

However, the goal of our model is not only to give a maturity score to a specific organization, but also provide them with some feedback and the necessary steps for reaching a higher maturity stage. For example, the overall maturity score for organization A is 2.72, which leaves a lot of room for improvement. Moreover, as the lowest score is for Data Modelling, a good starting point for higher maturity would be this category. Due to confidentiality reasons, more details regarding the maturity scores and feedback cannot be offered here.

Benchmarking – As already mentioned in the previous sections, the DWCMM can serve as a benchmarking tool for organizations. The DW maturity assessment questionnaire provides a quick way for organizations to assess their DW maturity and, at the same time, compare themselves in an objective way against others in the same industry or across industries. Of course, better results will be achieved for benchmarking after more organizations will take the maturity assessment. However, in order to have a better image on how the graph will look like when doing benchmarking, we will provide here an example for organization A using the data from the case studies we performed. The bar chart can be depicted below.

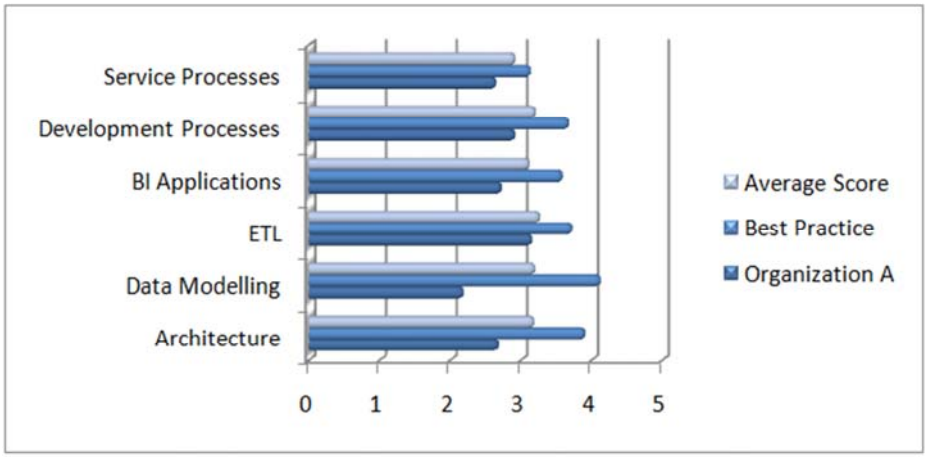

Figure 3: Benchmarking for Organization A.

To sum up, the DW maturity assessment questionnaire can be successfully applied in practice. We generally received positive feedback regarding the questionnaire from the case study interviewees. In this way, we could test whether the questions and their answers are representative for assessing the current DW solution for a specific organization and if they can be mapped to any organization depending on the situational factors. Respondents usually had no problems in recognizing the proposed benchmark categories and understanding the questions and answers from the survey. We also had the chance to apply the scoring method and give appropriate feedback for each case study. Finally, we combined all the feedback received from the case studies and did some minor, but valuable improvements to several questions and answers in order for them to be more representative for the analyzed characteristics and better fit the maturity stages.

7 Conclusions and Further Research

This research has been triggered by the estimates made by Gartner (2007) and other researchers that more than fifty percent of DW projects have limited acceptance or fail. Therefore, we developed a Data Warehouse Capability Maturity Model (DWCMM) that helps organizations assess the technical aspects of their current DW solution and provide guidelines for future improvements. This answered the main research question for our study: *How can the maturity of a company's data warehouse technical aspects be assessed and acted upon?*

The main conclusion from our research is that, even if our maturity model could help organizations improve their DW solutions, there is no "silver bullet" for a successful development of DW/BI solutions. The DWCMM provides a quick way for organizations to assess their DW/BI maturity and compare themselves in an objective way against others in the same industry or across industries. It received positive feedback from the five experts that reviewed and validated it and it also resonated well with the audiences from our four case studies. Several (mostly minor) improvements were made after the validation process.

Furthermore, emerging application domains relying on many diverse, mobile and heterogenenous data sources such as Mobile Learning & Analytics will benefit heavily from our DWCMM assessment tool to ensure a well-structured data foundation with a longitudinally prepared architecture.

However, our model is not without limitations. First of all, it is critical to emphasize the fact that the model only does a *high-level assessment*. In order to truly assess the maturity of their DW/BI solutions and discover the strong and weak variables, organizations should use our assessment as a starting point for a more thorough analysis. In the future, several questions could be added in our model for a more detailed analysis of the current DW/BI environment and more valuable feedback offered to organizations. Second, a limitation of this study is that it is based on the design science research which answers to research questions in the form of design artifacts. Being a qualitative research method, a risk for objectivity might arise.

Another limitation is related to the validation process for our model. Due to time constraints and difficulty of finding them, it was reviewed only by five experts. Therefore, more experts should be interviewed in the future to enrich the structure and content of the model. Also, due to the fact that the model was tested only in four cases, it is not possible to generalize the findings to any given similar situation. For further research, it would be interesting to validate the model using quantitative research methods. In this way, we will be able to do some statistical analysis on the data, more valuable benchmarking and improvements on the whole structure of the model. Another future extension that would increase the value of the model could include questions and analysis for other types of data modelling (*e.g.* normalized modelling, data vault, etc.) because, as stated earlier in this paper, we limited our maturity assessment only to dimensional modelling. Last, but not least, more work is also needed to extend our model to the analysis of DW/BI end user adoption and business value. New benchmark categories and maturity assessment questions regarding these two problems could also be added.

Acknowledgements

We would like to thank Inergy – and Frank Habers in particular – for sharing their experiences and providing prompt access to all the resources we needed for the research. We would also like to express our gratitude to all the experts and respondents who made room for us in their busy schedules and accepted to review our mode and fill in the assessment questionnaire, respectively.

References

AbuAli, A. & Abu-Addose, H., 2010. Data Warehouse Critical Success Factors. *European Journal of Scientific Research, 42,* pp. (2), 326-335.

April, A., Hayes, J., Abran, A. & Dumke, R., 2004. Software Maintenance Maturity Model: the Software Maintenance Process Model. *Journal of Software Maintenance and Evolution: Research and Practice, 17,* pp. (3), 197-223.

Arnott, D. & Pervan, G., 2005. A Critical Analysis of Decision Support Systems Research. *Journal of Information Technology, 20,* pp. (2), 67-87.

Azvine, B., 2005. Towards Real-Time Business Intelligence. *BT Technology Journal, 23,* pp. (3), 214-225.

Becker, J., Knacksedt, R. & Poppelbus, J., 2009. Developing Maturity Models for IT Management: A Procedure Model and its Application. *Business and Information Systems Engineering, 1,* pp. (3), 213-222.

Benbasat, I., Goldstein, D. & Mead, M., 1987. The Case Research Strategy in Studies of Information Systems. *MIS Quarterly, 11,* pp. (3), 369-386.

Cater-Steel, A., 2006. *Transforming IT Service Management - the ITIL Impact.* Adelaide, Australia, s.n.

Chauduri, S. & Dayal, U., 1997. An Overview of Data Warehousing and OLAP Technology. *ACM Sigmod Record, 26,* pp. (1), 65-74.

Colin, R., 2004. *An Introductory Overview of ITIL.* Reading, United Kingdom: itSMF Publications.

Darke, P., Shanks, G. & Broadbent, M., 1998. Successfully Completing Case Study Research: Combining Rigour, Relevance and Pragmatism. *Information Systems Journal, 8,* pp. (4), 273-289.

de Bruin, T., Freezey, R., Kulkarniz, U. & Rosemann, M., 2005. *Understanding the Main Phases of Developing a Maturity Assessment Model.* Sydney, Australia, s.n.

Gartner, 2007. *Creating Enterprise Leverage: The 2007 CIO Agenda.* [Online] [Accessed 24 June 2010].

Hayen, R., Rutashobya, C. & Vetter, D., 2007. An Investigation of the Factors Affecting Data Warehousing Success. *Issues in Information Systems, 8,* pp. (2), 547-553.

Hevner, A., March, S., Park, J. & Ram, S., 2004. Design Science in Information Systems Research. *Management Information Systems Quarterly, 28,* pp. (1), 75-106.

Hoffman, R., Shadbolt, N., Burton, A. & Klein, G., 1995. Eliciting Knowledge from Experts: A Methodological Analysis. *Organizational Behaviour and Human Decision Processes, 62,* pp. (2), 129-158.

Humphries, M., Hawkins, M. & Dy, M., 1999. *Data Warehousing: Architecture and Implementation.* New Jersey: Prentice Hall PTR.

Kaula, R., 2009. Business Rules for Data Warehouse. *International Journal of Information Technology, 5,* pp. 58-66.

Kaye, D., 1996. An Information Model of Organization. *Managing Information, 3,* pp. (6),19-21.

Kimball, R., 1996. *The Datawarehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses.* New York: John Wiley & Sons, Inc..

Kimball, R. & Caserta, J., 2004. *The Data Warehouse ETL Toolkit.* Indianapolis: Wiley Publishing, Inc..

Kimball, R. et al., 2008. *The Data Warehouse Lifecycle Toolkit, 2nd Edition.* Indianapolis: Wiley Publishing, Inc..

Klimko, G., 2001. *Knowledge Management and Maturity Models: Building Common Understanding.* Bled, Slovenia, s.n., pp. 269-278.

Lewis, J., 2001. *Project Planning, Scheduling and Control, 3rd Edition.* New York: McGraw-Hill.

March, S. & Hevner, A., 2007. Integrated Decision Support Systems: A Data Warehousing Perspective. *Decision Support Systems, 43,* pp. (3), 1031-1043.

Moss, L. & Atre, S., 2003. *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications.* Boston: Addison Wesley.

Nagabhushana, S., 2006. *Data Warehousing. OLAP and Data Mining..* New Delhi: New Age International Limited.

Navathe, S. B., 1992. Evolution of Data Modelling for Databases. *Communications of the ACM, 35,* pp. (9), 112-123.

Niessink, F. & van Vliet, H., 1999. *The IT Service Capability Maturity Model (IR-463),* Amsterdam: Division of Mathematics and Computing Science, Vrije Universiteit.

Niessink, F. & van Vliet, H., 2000. Software Maintenance from a Service Perspective. *Journal of Software Maintenance: Research and Practice, 12,* pp. (2), 102-120.

Paulk, M., Weber, C., Curtis, B. & Chrissis, M., 1995. *The Capability Maturity Model: Guidelines for Improving the Software Process.* Boston, MA: Addison-Wesley.

Ponniah, P., 2001. *Data Warehousing Fundamentals.* New York: John Wiley & Sons, Inc..

Sacu, C., Spruit, M. & Habers, F., 2010. *Data Warehouse (DW) Maturity Assessment Questionnaire (UU-CS-2010-021),* Utrecht: Utrecht University.

Salle, M., 2004. *IT Service Management and IT Governance: Review, Comparative.* [Online] Available at: http://www.hpl.hp.com/techreports/2004/HPL-2004-98.pdf [Accessed 16 July 2010].

Solomon, M., 2005. Ensuring a Successful Data Warehouse Initiative. *Information Systems Management, 22,* pp. (1), 26-36.

Vaishnavi, V. & Kuechler, W., 2008. *Design Science Research Methods and Patters: Innovating Information and Communication Technology.* Boca Raton, Florida: Auerbach Publications Taylor & Francis Group.

Vitt, E. & Luckevich, M., 2002. *Business Intelligence: Making Better Decisions Faster.*. Redmont: Microsoft Press.

Wixom, B. & Watson, H., 2001. An Empirical Investigation of the Factors Affecting Data Warehousing Success. *MIS Quarterly, 25,* p. (1).

Yin, R., 2009. *Case Study Research Design and Methods.* Thousand Oaks, California: SAGE Inc..

Appendix A: DWCMM Complete Maturity Matrix

DW TECHNICAL SOLUTION (28)				
ARCHITECTURE (8)				
INITIAL (1)	**REPEATABLE (2)**	**DEFINED (3)**	**MANAGED (4)**	**OPTIMIZED (5)**
Desktop data marts (*e.g.* Excel sheets)	Multiple independent data marts	Multiple independent data warehouses	A single, central DW with multiple data marts (Inmon) or conformed data marts (Kimball)	A virtual integrated DW
No business rules defined or implemented	Few business rules defined or implemented	Some business rules defined or implemented	Most business rules defined or implemented	All business rules defined or implemented
No metadata management	Non-integrated metadata by solution	Central metadata repository separated by tools	Central up-to-date metadata repository	Web-accessed central metadata repository with integrated, standardized, up-to-date metadata
No security implemented	Authentication security	Independent authorization for each tool	Role-level security at database level	Integrated companywide authorization security
CSVs files	Operational databases	ERP and CRM systems; XML files	Unstructured data sources (*e.g.* text or documents)	Various types of unstructured data sources (*e.g.* images, videos) and Web data sources
No methods to increase performance	Software performance tuning (*e.g.* index management, parallelizing and partitioning system, views materialization)	Hardware performance tuning (*e.g.* DW server)	Software and hardware tuning	DW specialized appliances
Desktop platform	Shared OLTP systems and DW environment	Separate OLTP systems and DW environment	Separate servers for OLTP systems, DW, ETL and BI applications	Specialized DW appliances (*e.g.* Netezza)
Monthly update or less often	Weekly update	Daily update	Inter-daily update	Real-time update
DATA MODELLING (8)				
INITIAL (1)	**REPEATABLE (2)**	**DEFINED (3)**	**MANAGED (4)**	**OPTIMIZED (5)**
No data modelling tool	Data modelling tools used only for design	Data modelling tools used also for maintenance	Standardized data modelling tool used for design	Standardized data modelling tool used also for maintaining metadata
No synchronization between data models	Manual synchronization of some of the data models	Manual or automatic synchronization depending on the data models	Automatic synchronization of most of the data models	Automatic synchronization of all of the data models

No differentiation between data models levels	Logical and physical levels designed for some data models	Logical and physical levels designed for all the data models	Conceptual level also designed for some data models	All data models have conceptual, logical and physical levels designed
No standards defined for data models	Solution-dependent standards defined for some of the data models	Enterprise-wide standards defined for some of the data models	Enterprise-wide standards defined for most of the data models	Enterprise-wide standards defined for all the data models
No documentation for any data models	Non standardized documentation for some of the data models	Standardized documentation for some of the data models	Standardized documentation for most of the data models	Standardized documentation for all the data models
Very few fact tables have their granularity at the lowest level possible	Few fact tables have their granularity at the lowest level possible	Some fact tables have their granularity at the lowest level possible	Most fact tables have their granularity at the lowest level possible	All fact tables have their granularity at the lowest level possible
No conformed dimensions	Conformed dimensions for few business processes	Conformed dimensions for some business processes	Enterprise-wide standardized conformed dimensions for most business processes; also making use of a high level design technique such as an enterprise bus matrix	Enterprise-wide standardized conformed dimensions for all business processes
Few dimensions designed; no hierarchies or surrogate keys designed	Some dimensions designed with surrogate keys and basic hierarchies	Most dimensions designed with surrogate keys and complex hierarchies	Slowly changing dimensions techniques (*i.e.*: type 2, 3 and more) also designed	Besides regular dimensions, special dimensions are also designed (*e.g.* mini, monster, junk dimensions)
ETL (6)				
INITIAL (1)	**REPEATABLE (2)**	**DEFINED (3)**	**MANAGED (4)**	**OPTIMIZED (5)**
Only hand-coded ETL	Hand-coded ETL and some standard scripts	ETL tool(s) for all the ETL design and generation	Standardized ETL tool and some standard scripts for better performance	Complete ETL generated from metadata
Simple ETL that just extracts and loads data into the data warehouse	Basic ETL with simple transformations such as: format changes, sorting, filtering, joining, deriving new calculated values, aggregation, etc and surrogate key generator	Advanced ETL capabilities: slowly changing dimensions manager, reusability, change data capture system, de-duplication and matching system, data quality system	More advanced ETL capabilities: error event table creation, audit dimension creation, late arriving data handler, hierarchy manager, special dimensions manager	Real-time ETL capabilities (optimization of ETL)
Daily automation: no; Specific data quality tools: no; Identifying data quality issues: no; Solving data quality issues: no	Daily automation: no; Specific data quality tools: no; Identifying data quality issues: yes; Solving data quality issues: no	Daily automation: no; Specific data quality tools: yes; Identifying data quality issues: yes; Solving data quality issues: no	Daily automation: yes; Specific data quality tools: yes; Identifying data quality issues: yes; Solving data quality issues: no	Daily automation: yes; Specific data quality tools: yes; Identifying data quality issues: yes; Solving data quality issues: yes
Restart and recovery system: no; Simple monitoring: no; Advanced monitoring: no; Real-time	Restart and recovery system: no; Simple monitoring: yes; Advanced monitoring: no; Real-time	Restart and recovery system: no; Simple monitoring: yes; Advanced monitoring: yes; Real-time	Restart and recovery system: yes; Simple monitoring: yes; Advanced monitoring: yes; Real-time	Restart and recovery system: yes; Simple monitoring: yes; Advanced monitoring: yes; Real-time

monitoring: no	monitoring: no	monitoring: no	monitoring: no	monitoring: yes
No standards	Few standards defined for ETL	Some standards defined for ETL	Most standards defined for ETL	All the standards defined for ETL
No metadata management	Business and technical metadata for some ETL	Business and technical metadata for all ETL	Process metadata is also managed for some ETL	All types of metadata are managed for all ETL
BI APPLICATIONS (6)				
INITIAL (1)	**REPEATABLE (2)**	**DEFINED (3)**	**MANAGED (4)**	**OPTIMIZED (5)**
Static and parameter-driven reports and query applications	Ad-hoc reporting; online analytical processing (OLAP)	Visualization techniques: dashboards and scorecards	Predictive analytics: data and text mining; alerts	Closed loop BI applications; real-time BI applications
BI tool related to the data mart	More than two tools for main stream BI (*i.e.*: reporting and visualization applications)	One tool recommended for main stream BI, but each department can use their own tool	One tool for main stream BI, but each department can use their own tool for specific BI applications (*e.g.* data mining, financial analysis, etc.)	One tool for main stream BI and one tool for specific BI applications
No standards	Few standards defined for BI applications	Some standards defined for BI applications	Most standards defined for BI applications	All the standards defined for BI applications
Objects defined for every BI application	Some reusable objects for similar BI applications	Some standard objects and templates for similar BI applications	Most similar BI applications use standard objects and templates	All similar BI applications use standard objects and templates
Reports are delivered manually on paper or by email	Reports are delivered automatically by email	Direct tool-based interface	A BI portal with basic functions: subscriptions , discussions forum, alerting	Highly interactive, business process oriented, up-to-date portal (no differentiation between operational and BI portals)
No metadata available	Some incomplete metadata documents that users ask for periodically	Complete up-to-date metadata documents sent to users periodically or available on the intranet	Metadata is always available through a metadata management tool, different from the BI tool	Complete integration of metadata with the BI applications (accessible through one button push on the attributes, etc.)
DW ORGANIZATION & PROCESSES (18)				
DEVELOPMENT PROCESSES (10)				
INITIAL (1)	**REPEATABLE (2)**	**DEFINED (3)**	**MANAGED (4)**	**OPTIMIZED (5)**
ad-hoc development processes; no clearly defined development phases (*i.e.*: planning, requirements definition, design, construction, deployment, maintenance)	repeatable development processes based on experience with similar projects; some development phases clearly separated	standard documented development processes; iterative and incremental development processes with all the development phases clearly separated	development processes continuously measured against well-defined and consistent goals	continuous development process improvement by identifying weaknesses and strengthen the process proactively, with the goal of preventing the occurrence of defects
no separation between environments	two separate environments (*i.e.*: usually development	some separation between environments (*i.e.*: at least three	some separation between environments (*i.e.*: at least three	all the environments are distinct with automatic transfer between them

	and production) with manual transfer between them	environments) with manual transfer between them	environments) with automatic transfer between them	
no standards defined	few standards defined	some standards defined	a lot of the standards defined	a comprehensive set of standards defined
no quality assurance activities	ad-hoc quality assurance activities	standardized and documented quality assurance activities done for all the development phases	level 3) + measurable and prioritized goals for managing the DW quality (*e.g.* functionality, reliability, maintainability, usability)	levels 4) + causal analysis meetings to identify common defect causes and subsequent elimination of these causes; service quality management certification
no project sponsor	IT project manager	chief information officer (CIO) or an IT director	single sponsor from a business unit or department	multiple individual sponsors from multiple business units or departments
no project management activities	project planning and scheduling	some of the main project management activities (project planning and scheduling; project risk management; project tracking and control)	some project management activities; standard and efficient procedure and documentation	project planning and scheduling; project risk management; project tracking and control; standard and efficient procedure and documentation; evaluation and assessment
no formal roles defined	defined roles, but not technically implemented	formalized and implemented roles and responsibilities	level 3) + periodic peer reviews (*i.e.*: review of each other's work)	level 4) + periodic evaluation and assessment of roles (*i.e.*: assess the performance of the roles and match the needed roles with responsibilities and tasks)
ad-hoc knowledge gathering and sharing	organized knowledge sharing through written documentation and technology (*e.g.* knowledge databases, intranets, wikis, etc.), and also through training and mentoring programs	knowledge management is important to top level management; knowledge creation and sharing through brainstorming, training and mentoring programs	central business unit knowledge management; quantitative knowledge management control and periodic knowledge gap analysis	continuously improving inter-organizational knowledge sharing
ad-hoc requirements definition; no methodology used	methodologies differ from project to project; interviews with business users for collecting the requirements	standard methodology for all the projects; interviews and group sessions with both business and IT users for collecting the requirements	level 3) + qualitative assessment and measurement of the phase; requirements document also published	level 4) + causal analysis meetings to identify common bottlenecks causes and subsequent elimination of these causes
only unit testing is done; no standards or	other types of testing are beginning to be done	diverse types of testing; some standards	diverse types of testing; standard procedure and	all the main types testing (unit testing by

documentation	(some of the following: unit testing by another person; system integration testing; regression testing; acceptance testing)		documentation	another person; system integration testing; regression testing; acceptance testing); user training; standard procedure and documentation; external assessments and reviews
SERVICE PROCESSES (8)				
INITIAL (1)	**REPEATABLE (2)**	**DEFINED (3)**	**MANAGED (4)**	**OPTIMIZED (5)**
reactive service quality management	ad-hoc service quality management	proactive service quality management including a standard procedure	level 3) + service quality measurements periodically compared to the established goals to determine the deviations and their causes	levels 4) + causal analysis meetings to identify common defect causes and subsequent elimination of these causes; service quality management certification
ad-hoc knowledge gathering and sharing	organized knowledge sharing through written documentation and technology (*e.g.* knowledge databases, intranets, wikis, etc.), and through training and mentoring programs	knowledge management is important to top level management; knowledge creation and sharing through brainstorming, training and mentoring programs	central business unit knowledge management; quantitative knowledge management control and periodic knowledge gap analysis	continuously improving inter-organizational knowledge sharing
customer service needs documented in an ad-hoc manner; no service catalogue compiled	some customer service needs documented and formalized	all the customer service needs documented and formalized according to a standard procedure into service level agreements (SLAs)	SLAs reviewed with the customer on both a periodic and event-driven basis	actual service delivery continuously monitored and evaluated with the customer on both a periodic and event-driven basis for continuous improvement (SLAs including penalties)
incident management is done ad-hoc with no specialized ticket handling system or service desk to assess and classify them prior to referring them to a specialist	a ticket handling system is used for incident management	a service desk is the recognized point of contact for all the customer queries; incidents assessment and classification is done following a standard procedure	standard reports concerning the incident status including measurements and goals (*e.g.* response time) are regularly produced for all the involved teams & customers; an incident management database is established as a repository for the event records	trend analysis in incident occurrence and also in customer satisfaction and value perception of the services provided to them
change requests are made and solved in an ad-hoc manner	a change management system is used for storing the requests for	a standard procedure is used for approving, verifying, prioritizing	standard reports concerning the change status including	trend analysis and statistics regarding change occurrence,

	change	and scheduling changes	measurements and goals (*e.g.* response time) are regularly produced for all involved teams&customers; standards established for documenting changes	success rate, customer satisfaction and value perception of the services provided to them
no resource management activities established within the organization	ad-hoc resource management activities (only when there is a problem)	resource management is done constantly following a standard procedure	standard reports concerning performance and resource management including measurements and goals are done on a regular basis	resource management trend analysis and monitoring to make sure that there is sufficient capacity to support planned services
no availability management (reactive availability management)	ad-hoc availability management	availability management documented and done using a standardized procedure (all elements are monitored	risk assessment to determine the critical elements and possible problems	availability management trend analysis and planning to ensure that all elements are available for the agreed service level targets
ad-hoc changes solving and implementation; no release naming and numbering conventions	release naming and numbering conventions	release management is documented and done following a standardized procedure; assigned release management roles and responsibilities	standard reports concerning release management including measurements and goals are done on a regular basis; master copies of all software in a release secured in a release database	release management trend analysis, statistics and planning