An Environment and Approach for the Evaluation of Database Design Processes

Juan Garbajosa, Pedro P. Alarcon, Hector Garcia, Maria Alandes, Ivan Bernardo, Carolina Fernandez

Technical University of Madrid. E.U. Informatica. Dpt. OEI. Ctra. Valencia. Km.7, E-28031 Madrid, Spain. e-mail: (jgs, pcavero, hgarcia)@eui.upm.es, malandes@bicho.eui.upm.es

Mario Piattini

University of Castilla – La Mancha (UCLM). E.U. Informatica. Ronda de Calatrava 5, E-13004 Ciudad Real, Spain. E-mail: mpiattin@inf-cr.uclm.es

Abstract

Conducting a process assessment is a difficult task that requires to take various issues under consideration. Whilst general (development, procurement) processes have been widely studied and multiple references in literature exist, specific processes (e.g. database design) have not received so much attention. In order to face this problem a study on the items related to database design process evaluation has been performed. SPICE/ISO-IEC 15504 has become a popular model to perform assessments, and it is internationally accepted. For this reason it has been used as our reference model. The results of our work, following the specifications defined in the process dimension of SPICE are described within this paper.

In order to alleviate the task of managing the assessment, we have developed an integrated tool that provides assistance in the evaluation processes. It offers a friendly and easy-to-use interface to carry out assessments. This tool intends to support to the assessor in the task of deciding the level of achievement of the work products and base practices. This tool generates reports and statistics from the evaluation information. It also provides with the possibility of having more than one instance of evaluation performed by different assessors. Tool basic guidelines are presented.

1. Introduction

Nowadays databases play an essential role in systems such as network management, data warehouse and command systems. Good performance of the systems is fundamental and relies on an appropriate database with a *good* design. The complexity of these systems and their associated databases is increasing at a high speed. Therefore database design process assessments play an important role in the task of guaranteeing that the new functionalities of databases will be attained.

Current process assessments models must be general enough in order to be applied, in principle, to any kind of software process instance. Therefore, it might happen that, for some specific processes, existing models could not deal with all the information required to properly assess the process [1]. We understand that this is the case for database design process and SPICE/ISO-IEC 15504 [4].

SPICE/ISO-IEC 15504 is now widely accepted as a reference model for process assessments and this has been one of the most important reasons to take it as our framework. Then we studied the database design process and obtained a mapping of a *generic database design process* onto the SPICE/ISO-IEC 15504 process dimension. A summarized description of the results, originally described in [2,3], is presented in this article.

As we believe that the task of conducting an assessment is hard and deals with a lot of information, we thought that it could be of interest to implement a tool that could give support during the assessment of database design processes. The tool stores intermediate results and generates reports. Other tools are described in literature [5]. Our tool is specific to the process of defining database schemas. An overview of this tool is described in the article. To end the paper a number of conclusions are presented.

2. Database design process assessment

As we mentioned in the introduction EMEDITA was the starting point to map a generic database design process onto SPICE/ISO-IEC 15504.

As figure 1 shows, EMEDITA is comprised of a set of phases, which contain stages which, in their turn, contain tasks and finally, basic activities, subtasks [7]. Subtasks are taken into account to develop the mapping with SPICE base practices. EMEDITA has been developed according to METRICA framework, the Spanish government system-development methodology [11]. EMEDITA is compliant with ISO/IEC 12207 [10] and has considered database design usual practices as described in [6,8,9]. The findings of mapping a Generic database design process onto the SPICE-ISO/IEC 15504-process model were described in [2]. The most relevant outputs are summarized below, and some of the basics of the mapping guidelines are presented in here as well.

The Capability dimension has been analyzed and, basically, no changes have been introduced except for those coming indirectly from the process dimension.

SPICE is defined on a general-purpose basis. We studied up to what extent SPICE-ISO/IEC 15504 offers guidance for database design process assessment and improvement. Specific characteristics of database become part of a problem when they are treated like any other conventional software module such a control module or a man machine interface.

In order to analyze the appropriateness of SPICE for database developments, first, mapping between a generic database design process, obtained from EMEDITA, and SPICE was carried out. A group of tasks that, eventually, could be mapped onto SPICE base practices were identified. The second step was to define a number base practices that we considered adequate to support a database assessment.

Only the process dimension was taken into consideration as a starting point. The mapping was done maintaining the maximum coherence degree between both models and trying to adjust as much as possible their contents. Mapping showed that database design tasks are not fully contemplated in SPICE. We could say that database design tasks would not feel at ease in SPICE. Once processes were mapped, the second step was to find a group of base practices where they could fit, as maintaining both models philosophy.

The problem of comparing conventional software concepts with database concepts raised up in the base practices mapping. The architectural design concept was considered to be similar to the idea of he conceptual schema definition in database, as well as the concept of detailed design was compared with the rest of subtasks performed to achieve a database physical design. The mapping was carried out taking this criteria into consideration. Figure 2 shows the result.

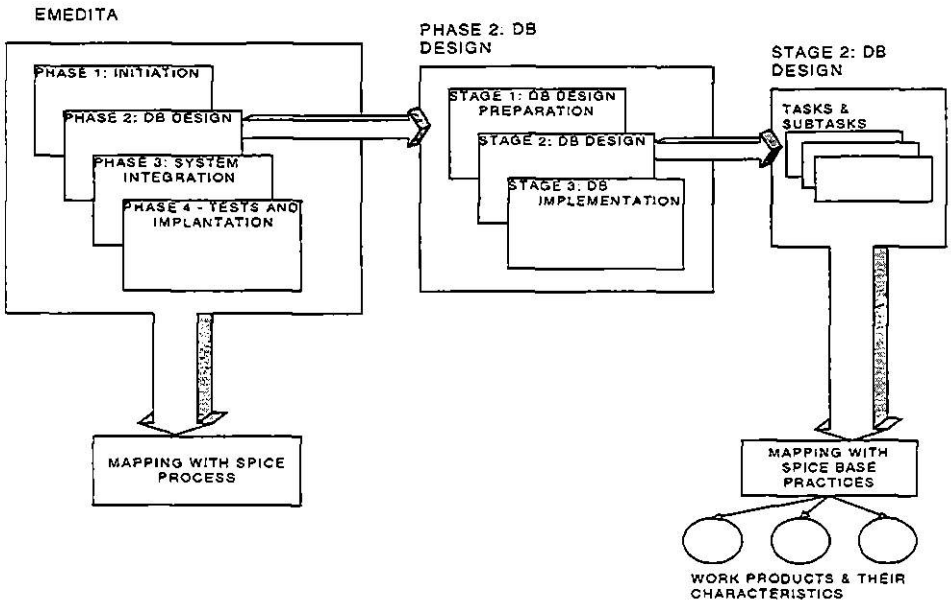

Fig. 1. EMEDITA structure and mapping philosophy

Finally the work products of SPICE base practices were analyzed from the database point of view. They represented the most difficult part of the mapping as some of SPICE work products matched those of the database design methodology, but their needed characteristics in a database context were rather different from the SPICE specifications. Some other database work products were missing in SPICE.

The first conclusion that we can draw is that the mapping between the database design methodology and the Spice reference model is not that easy. Actually we understand that database design tasks have not been fully contemplated within SPICE. We could say that database design tasks would not feel at ease in SPICE.

EMEDITA Processes	DB	DW	15504	15504 Processes
Project Initiation	X	X		
Project planning & goals definition	X	X	CUS.1	Acquisition process
Project goals identification	X	X	CUS.1.1	Acquisition preparation process
Involved areas identification	X	X	CUS.1.1	Acquisition preparation process
Participants identification & organisation	X	X	ORG.3	Human resource management process
Project general planning	X	X	MAN.2	Project management process
Project start	X	X	CUS.1.1	Acquisition preparation process
Project scope definition	X	X	CUS.1.1	Acquisition preparation process
Involved areas study	X	X	CUS.1.1	Acquisition preparation process
Involved areas data needs	X	X	CUS.1.1	Acquisition preparation process
Project environment definition	X	X	CUS.2	Supply process
Initial planning preparation	X	X	MAN.2	Project management process
Cost-benefit analysis	X	X	CUS.1.1	Acquisition preparation process
Revisit contents	X	X	CUS.1.2	Supplier selection process
Initial planning refinement	X	X	CUS.1.2	Supplier selection process
Standards adoption	X	X	CUS.1.2	Supplier selection process
Database design	X			
Data warehouse design		X		
Database design preparation	X		CUS.1	Development process
Data warehouse design preparation		X	CUS.1	Development process
Data source study	X	X	ENG.1.1	System requirements analysis & design process
Event analysis		X	ENG.1.1	System requirements analysis & design process
Storage capability estimation	X	X	ENG.1.1	System requirements analysis & design process
Database design	X			
Data warehouse design		X		
System architecture definition		X	ENG.1.1	System requirements analysis & design process
Logical model design	X	X	ENG.1.3	Software design process
Physical model design	X	X	ENG.1.3	Software design process
Data source analysis	X	X	ENG.1.5	Software requirements analysis process
Interface design	X	X	ENG.1.5	Software requirements analysis process
Architecture components analysis		X	ENG.1.1	System requirements analysis & design process
Revisit component specification		X	CUS.2	Supply process
Standard packages evaluation	X	X	CUS.1.1	Acquisition preparation process

Data quality strategies definition	X	X	MAN.3	Quality management process
Inherited components integration study	X	X	ENG.1.1	System requirements analysis & design process
System initial documentation	X	X	SUP.1	Documentation process
Architecture components implementation		X	ENG.1.4	Software construction process
System integration study	X	X		
System integration planning	X	X		
Participants identification & organisation	X	X	ORG.3	Human resource management process
General planning	X	X	MAN.2	Project management process
Assembly cost estimation	X	X	ENG.1.5	Software integration process
Physical model refinement	X	X	ENG.1.5	Software integration process
Assembly environment preparation	X	X	ENG.1.5	Software integration process
Test plan development	X	X	ENG.1.3	Software design process
Data loading	X	X	ENG.1.5	Software integration process
Final documentation generation	X	X	SUP.1	Documentation process
System integration	X	X	ENG.1.5	Software integration process
Tests	X	X	ENG.1.6	Software testing process
Data load test	X	X	ENG.1.6	Software testing process
System test	X	X	ENG.1.6	Software testing process
Tests verification	X	X	ENG.1.6	Software testing process
System implantation & acceptance test	X	X		
System implantation planning	X	X		
User training plan developement	X	X	ORG.3	Human resource management process
Implantation environment stablishment	X	X	ENG.1.7	System integration & testing process
Final data loading	X	X	ENG.1.7	System integration & testing process
System implantation	X	X	ENG.1.7	System integration & testing process
Hardware & software systems adaptation	X	X	ENG.1.7	System integration & testing process
Existing data conversion	X	X	ENG.1.7	System integration & testing process
Current system stop & backup	X	X	ENG.1.7	System integration & testing process
Conversion rutines execution	X	X	ENG.1.7	System integration & testing process
Acceptance testing	X	X	ENG.1.7	System integration & testing process
Acceptance test preparation	X	X	ENG.1.7	System integration & testing process
Acceptance test execution	X	X	ENG.1.7	System integration & testing process
Last chages incorporation	X	X	ENG.1.7	System integration & testing process
‡ Process not common to DB & DW				¥ It can be mapped to different processes

DB: Relational Database DW: Datawarehouse

Fig. 2. EMEDITA vs SPICE

As far as base practices are concerned, the problem of concepts that cannot be assimilated raises up. In the mapping between SPICE base practice Develop software architecture and all the practices of the first design approach, there was a need to compare the concept architecture to conceptual design. It seems that it is possible to integrate both ideas. However the uncertainty still remains, because this mapping is forced by the need of expressing the methodology with SPICE, overlooking the lack of accuracy in this decision.

This inflexibility is a consequence of the existing gap between database design and conventional software modules design that, from our point of view, SPICE does not address with enough depth.

Work products have represented the most difficult part of the mapping. Some of them turned out to be valid for the database methodology, but the characteristics were completely different from SPICE specifications, which means that there is no criteria to follow in SPICE to judge whether a database product has been successfully produced or not.

Some other database work products did not appear at all in SPICE as the base practices that produce them were also missing. An important case was Data Quality, a difficult practice for which SPICE offers no guidance to perform an assessment. We are also studying this issue more in depth and some first results were presented in [profess].

3. An integrated support tool

The main goal of the tool is to provide the assessor with support for evaluating database processes. In order to carry out the evaluation EMEDITA was used with the modifications that are explained above. Both *process* and *capability dimension* have been implemented. Capability dimension has been maintained basically unchanged with respect to the SPICE model.

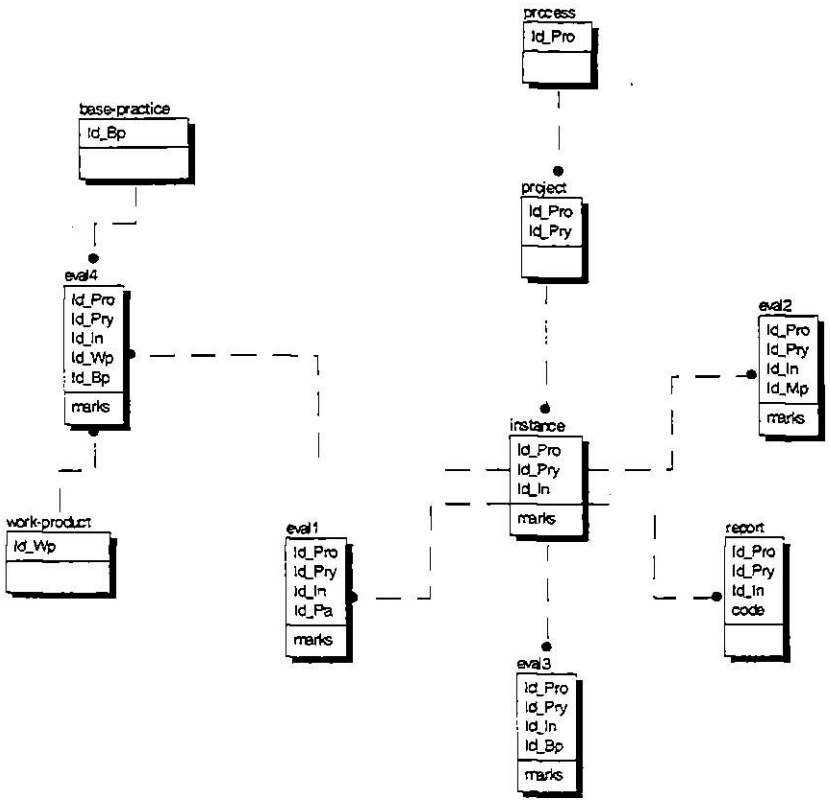

Fig. 3 – Database schema

The tool is comprised by a graphical user interface that basically displays information related to EMEDITA (work products, base practices, management practices, process attributes...) and a database that stores the results of the evaluation. The tool database is accessed through ODBC, which adds flexibility to migrate to a different database management system.

The data model schema is presented in figure 3. Within this model each evaluation is considered as a project within the tool. It is possible to have different projects for the same set of processes. Within a project it is also possible to carry out several assessments, that can be seen as snapshots of the set of processes. Each snapshot is called *instance*. One instance is comprised of an evaluation mark and a report of such evaluation. Several base practices and work products can be analyzed in an evaluation.

A different evaluator can produce each instance. The evaluation goals can be different, or they can be just performed in different moments. Thus, a project may have several instances, each of them containing specific information of the circumstances under which the evaluation was performed. This information is besides fundamental to correctly interpret the results of any evaluation. Whenever an evaluation is performed, a project and an instance must be specified.

The *Process Dimension* evaluation offers the list of *Base practices* and *Work products* of all the database design processes. The *Base practices* and *Work products* are then measured according to the evaluator's criteria and the given values are stored.

The *Capability Dimension* evaluation offers the list of *Process Attributes* and *Management Practices* associated to each database design process, and the evaluator also estimates them. The values of each evaluated process are stored in the database

In the following paragraphs we present an example, and we explain the evaluation procedure for the process and capability dimensions.

We have considered the subtask 2.2.1.1 *Entity Identification* in EMEDITA and only one base practice, BP1, *Determine entities and their description* as the following synoptic diagram represents:

Phase 2 : database design
↳ Stage 2.2: database design
↳ Subtask 2.2.1.1. Entity Identification
↳ Practice 2.2.1.1.BP1: Determine entities and their description.

Work products:
Input: SW requirements
Output: High Level SW design

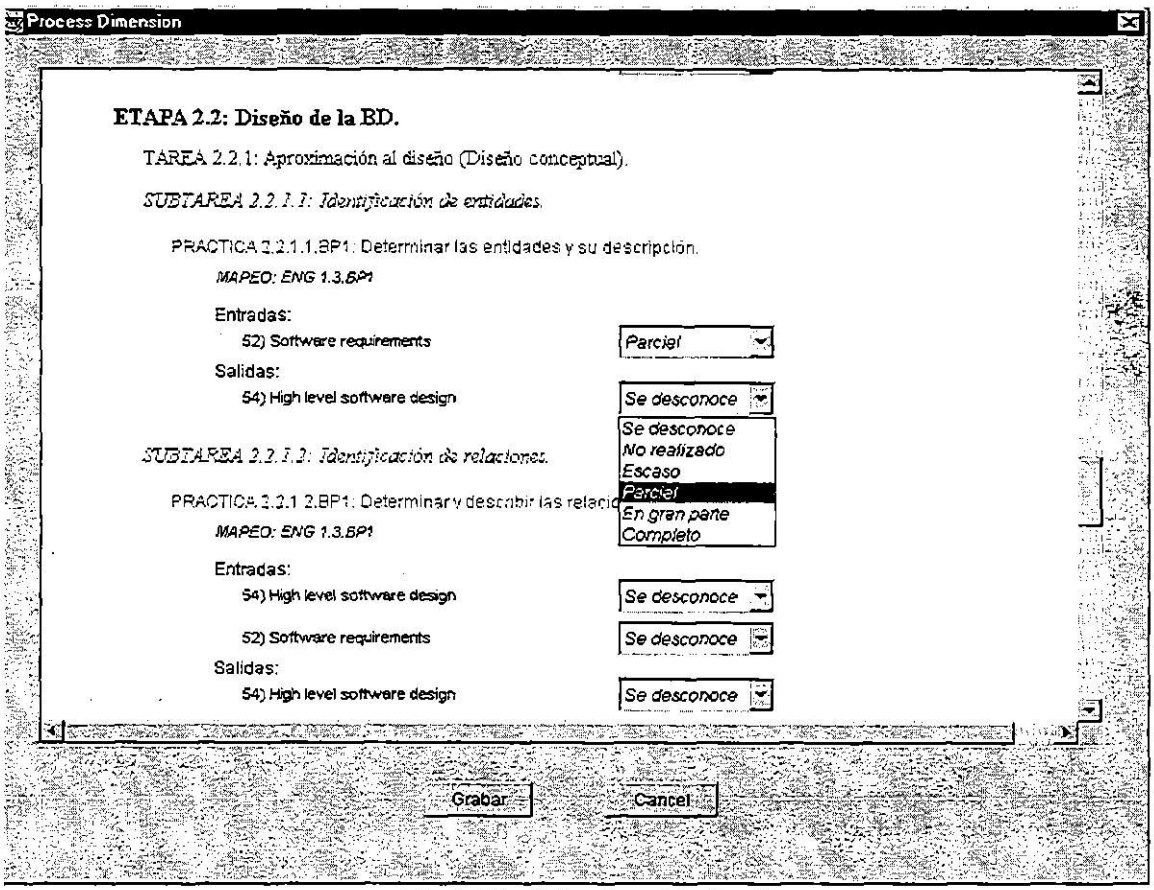

Fig 4- Work Products Evaluation

3.1. Process dimension evaluation

This subtask has two work products: *Software requirements*, which is an input work product and *High level software design*, which is an output workproduct.

In the evaluation we first evaluate the workproducts. The evaluator assigns a value for these two workproducts according to his criteria. The values can be: Fully, Largelly, Partially, None and Unknown as shown in Figure 4.

When the work products are evaluated we can assess the base practices. The associated work products are displayed in order to give supporting evidence to the achievement rating that will be assigned to the base practice. The rating values are also: Fully, Largelly, Partially, None and Unknown.

When the base practice is evaluated, we have completed the process dimension evaluation and a report can be then generated with the assigned ratings.

3.2. Capability dimension evaluation

For our process all the Management Practices are displayed and a value of Fully, Largelly, Partially, None and Unknown can be assigned to them. Once the management practices are evaluated we can continue on evaluating the Process Attributes. Fully, Largelly, Partially, None and Unknown are the values that can be assigned to them.

Finally we can generate a report of the Capability Dimension Evaluation.

4. Conclusions

Within this paper we have described the results of a study performed with the objective of analyzing the validity of model assessments, such as ISO/IEC 15504 SPICE, for specific processes such as database design. We have also presented a tool specifically design to support the evaluation of data base design processes.

After studying the impact of database design processes in current frameworks, such as SPICE/ISO-IEC 15504, our conclusion was that SPICE/ISO-IEC 15504 is oriented towards general software processes and, therefore, there exist a lack of specificity. SPICE has proved as a powerful framework but lacking some specific issues such as some base practices and workproducts. The evaluation tool that has been produced in the context of the project takes into account all those issues.

5. Acknowledgements

The work presented within this paper has been performed in the context of the project PROCBD and under a contract with CRONOS IBERICA S.A. This project is partially supported by the Ministry of Industry of Spain Ref. T87/1998, Programme ATYCA.

6. References

[1]. *Data Quality and Database Design Process Assessment* - Juan Garbajosa, Pedro P. Alarcon, Hector Garcia, Maria Alandes, Macario Polo - DIQ2000. Orlando, 2000.

[2]. *Database design in the context of the SPICE process dimension* - Juan Garbajosa, Hector Garcia, Maria Alandes, Mario Piattini - 1st International SPICE conference, SPICE'2000,Limmerick, 2000.

[3]. *Introducing the data role in models for database assessment* - Juan Garbajosa, Pedro P. Alarcon, Hector Garcia, Maria Alandes, Mario Piattini - International Conference on Product Focused Software Process Improvement, PROFESS'2000,Oulu, 2000.

[4]. *ISO/IEC FTCI TR 15504*: 1999 Information Technology – Software Process Assessment. Parts 1-9.

[5]. *The SEAL Process Assessment Tool* – 1999 Software Engineering Applications Laboratory. South Africa.

[6]. *Database Design. An Entity Relationship Approach* - Batini, C., Ceri, S., Navathe, S. Benjamin Cummings, 1992

[7]. *Methodology for the Development of Data Warehouse Projects (in Spanish)* - Carmona, O., García H. - Minor Thesis. Technical University of Madrid. E.U. Informatica. 1999.

[8]. *Database Models (In Spanish)* - De Miguel, A., Piattini, M. - RA-MA. 1999.

[9]. *Fundamentals of Database systems* - Elsmarsi, R. , Navathe, S. - Addison Wesley, 1997

[10]. *ISO/IEC JTC1 1 12207*:1995 Software Lifecycle Processes

[11]. *Methodology for Planning and Development of Information systems METRICA version 2*. Ministry of Public Administration of Spain.