Programme Committee

Vicki Almstrum
University of Texas, USA
Mike Berry
CISRO, Australia
Jorgen Boegh
Delta, Denmark
Pearl Brereton
Keele University, United Kingdom
David Budgen
Keele University, United Kingdom
Peter Comer
University of Sunderland, United Kingdom
Ray Dawson
Loughborough University, United Kingdom
Stefano de Panfilis
Engineering-Ingeniria Informatica SpA, Italy
Khaled El Emam
National Research Council of Canada
Rachel Harrison
University of Reading, United Kingdom
Martin Host
Lund University, Sweden
Magne Jorgensen
University of Oslo, Norway
Barbara Kitchenham
Keele University, United Kingdom
Stephen Linkman
Keele University, United Kingdom
Emelia Mendes
University of Auckland, New Zealand
Alberto Pasquini
ENEA, Italy
Shari Lawrence Pfleeger
Systems/Software Inc. USA
Lesley Pickard
Keele University, United Kingdom
Martin Sheppherd
Bournemouth University, United Kingdom
Daj Sjoberg
University of Oslo, Norway
Walter Tichy
Karlsruhe University, Germany
Claes Wohlin
Blekinge Institute of Technology, Sweden

Sponsorship

EASE is hosted by

The Department of Computer Science
Keele University

EASE is supported by
the following organisations:

CSR
Centre for **S**oftware **R**eliability

BCS
British **C**omputer **S**ociety

IEE
Institute of **E**lectrical **E**ngineers

Keynote speakers will be

Professor Janice Singer of the National Research Council of Canada is the opening keynote speaker.

Professor Barbara Kitchenham of Keele University is the closing keynote speaker.

Designed and Produced by:
Keele Graphic Services: 01782 584215 or 583092

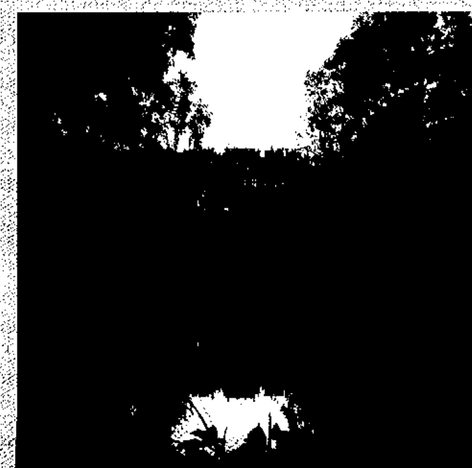
EASE

Evaluation and Assessment in Software Engineering

at

KEELE
UNIVERSITY

8th to 10th April 2002



EASE Programme 2002

Monday April 8th

Time	Title	Authors
11:00	Welcome	
11:15	**Keynote:** So you're thinking of doing an experiment: Ethics and how they play a role in EASE	Janice Singer
12:30	Lunch	
13:30	**Session 1**	
	Case Study: Evaluating the effect of E-mail interruptions within the workplace	Thomas Jackson, Ray Dawson Darren Wilson
	Four Degrees of Separation: Intrusive and non-intrusive data collection	Stephen Owen, David Budgen Pearl Brereton
	On the vital role and difficulty of definition and measurement for the validation of S.E. theory	Franck Xia
15:00	Coffee	
15:30	**Session 2**	
	Supporting Communicability with Use Case Guidelines: An Empirical Study	Keith Phalp, Karl Cox
	Comparing effort estimates based on Use Case Points with expert estimates	Bente Anda
	Requirements Problems in Twelve Software Companies: An Empirical Analysis	Tracey Hall, Sarah Beecham, Austen Rainer
	Techniques for gathering information about an organisation and its needs for software process improvement	Espen Koren, Hans Westerheim

Tuesday April 9th

Time	Title	Authors
09:00	**Keynote** Practical Aspects of Bringing Mathematical Innovations & Software Applications to Business	Glenn Koller
10:00	**Session 3**	
	A Framework for Evaluating a Software Bidding Model	Barbara Kitchenham, Lesley Pickard, Peter Jones, Stephen Linkman
10:45	Coffee	
11:15	**Session 4**	
	Benchmarking of Processes for Managing Product platforms - a Case Study	Martin Höst, Enrico Johansson, Adam Noren, Lars Bratthall
	An approach towards a rapid universal software evaluation scheme	Franz Gruber, Christa Illibauer
	A family of experiments to investigate the influence of context on the effect of inspection techniques	Marcus Ciolowski, Forrest Shull, Stefan Biffl
13:00	Lunch	
14:00	**Session 5**	
	Experiment about Test-first programming	Matthias Müller, Oliver Hagner
	Investigating the Influence of Software Inspection process Parameters on Inspection Meeting Performance	Michael Halling, Stefan Biffl
15:30	Coffee	
	Session 6	
16:00	Aggregating Viewpoints for Strategic Software Process Improvement – a Method and Case Study	Daniel Karlstrom, Per Runeson, Claes Wohlin
	Combining quantitative software development cost estimation precision data with qualitative data from project experience reports at Ericsson design centre in Norway	Magne Jorgensen, Norman Lovstad, Liv Moen
	Investigation of product process dependency models through probabilistic modelling	Manoranjan Satpathy, Rachel Harrison, Daniel Rodriguez

Wednesday April 10th

Time	Title	Authors
09:00	**Session 7**	
	Validating Metrics for Data Warehouses	Manuel Serrano, Coral Calero, Mario Piattini
	An empirical study of software development across time zones.	Adel Taweel, Pearl Brereton
	Business needs driving IT decisions: using feature analysis and stakeholder evaluation in Rolls-Royce	Mark de Chazal, Heulwen Pearce, Ray Dawson
10:45	Coffee	
11:15	**Session 8**	
	Making Inferences with Small Numbers of Training Sets	Colin Kirsopp, Martin Shepperd
12:00	**Keynote** Five Things I Hate About Empirical Software Engineering	Barbara Kitchenham
	Closing Remarks	David Budgen
13:15	Lunch	

EASE 2002

Proceedings of the Conference on Empirical Assessment in Software Engineering

At Keele University
8th. – 10th. April 2002
Sponsored by
The Institute of Electrical Engineers
The British Computer Society
The Centre for Software Reliability
and is supported by
Keele University

Contents

Validating Metrics for Data Warehouses

Manuel Serrano, Coral Calero, Mario Piattini,
{manuel.serrano, coral.calero, mario.piattini}@uclm.es
Alarcos Research Group
Escuela Superior de Informática
University of Castilla – La Mancha
Paseo de la Universidad, 4
13071 Ciudad Real
Spain

Abstract

A data warehouse (DW) is a set of data and technologies aimed at enabling the executives, managers and analyst to make better and faster decisions. So organizations are adopting data warehouses to manage information efficiently as "the" main organizational asset. Due to the principal role of data warehouse in taking strategic decisions its quality is fundamental. The design of the star schema of a data warehouse is one of the most important factors affecting the quality of the final system. In the last years different authors have proposed some useful guidelines to design to design data warehouse models, however more objective indicators are needed to help designers and managers to develop quality data warehouses. In this paper we will propose a set of metrics for data warehouse models and we will show their empirical validation.

Keywords: data warehouse quality, data warehouse metrics.

1. Introduction

Organizations are very rich in data but poor in information. Nowadays organizations can store vast amounts of data obtained at a relatively low cost, however these data fail to provide information [11]. Data warehouses have appeared as a solution to this problem supporting decision making processes and new kind of applications as marketing.

A data warehouse is defined as a "collection of subject-oriented, integrated, non-volatile data that supports the management decision process" [13]. Data warehouses have become the key trend in corporate computing in the last years, since they provide managers with the most accurate and relevant information to improve strategic decisions. Jarke et al. [14] forecast a 12 Millions American dollars for the data warehouse market. However, the success of implement a data warehouse and its use in the organizations could be seriously affected by the lack of quality.

Recently some methodologies for data warehouse design have been proposed [12] [15] [16], but they are not enough to assure data warehouse quality. This lack of quality could be at two levels: data models and data of the data warehouse.

Our work focuses on dimensional data model quality. Different authors have suggested interesting recommendations for achieving a "good" dimensional data model [16] [1] [13]. However, quality criteria are not enough on their own to ensure quality in practice, because different people will generally have different interpretations of the same concept. According to the Total Quality Management (TQM) literature, measurable criteria for assessing quality are necessary to avoid "arguments of style" [5]. The objective should be to replace intuitive notions of design "quality" with formal, quantitative measures in order to reduce subjectivity and bias in the evaluation process. However, for data modelling to progress from a "craft" to an engineering discipline, the desirable qualities of data models need to be made explicit [17].

The final aim of our work is to define a set of metrics for assuring data warehouse quality by means of measuring the dimensional data model quality. In the next section we will present the method we use for defining correct metrics. A proposal of metrics for data warehouses will be described in section 3 and an example of the proposed metrics will be shown in section 4. Sections 5 will present the empirical validation of the metrics and conclusions and future work will come in the last section.

2. Defining Valid Metrics

Metrics definition must be done in a methodological way and it is necessary to follow a number of steps to ensure the reliability of the proposed metrics. Figure 1 presents the method we follow for the metrics proposal [8].

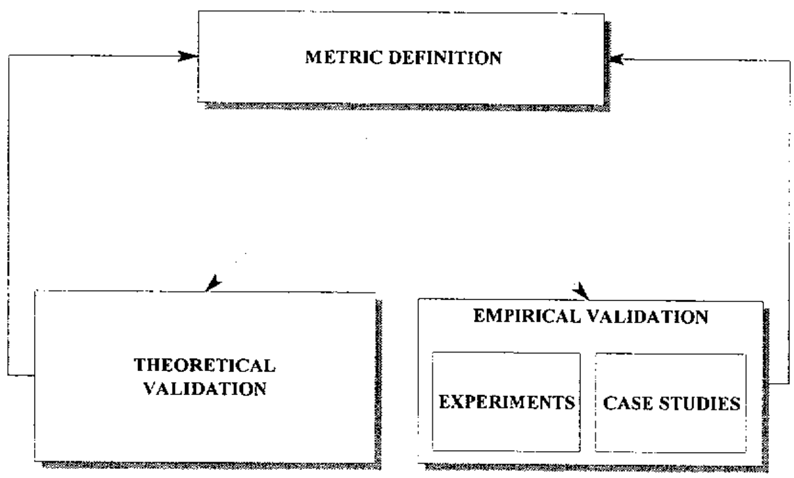


Figure 1. Steps followed in the definition and validation of metrics

In this figure we have three main activities:

- **Metrics definition.** The first step is the proposal of metrics. This definition is made taking into account the specific characteristics of the system we want to measure and the experience of designers of these systems. A goal-oriented approach, as GQM (Goal-Question-Metric) [2] can also be very useful in this step.

- **Theoretical validation.** The second step is the formal validation of the metrics. The formal validation helps us to know when and how apply the metrics. There are two main tendencies in metrics validation: the frameworks based on axiomatic approaches [19] [7] and the ones based on the measurement theory [20] [22]. The strength of measurement theory is the formulation of empirical conditions from which we can derive hypothesis of reality. The final information when applying this kind of frameworks in to know to which scale a metric pertains and based on this information we can know which statistics and which transformations can be done with the metric[22].
- **Empirical validation.** The goal of this step is to prove the practical utility of the proposed metrics. Although there are various ways of performing this step, basically we can divide the empirical validation into experimentation and case studies [4] [10].

As shown in figure 1, the process of defining and validating metrics is evolutionary and iterative. As a result of the feedback, metrics could be redefined based on discarded theoretical or empirical validations.

3. Metrics for Datawarehouses

Taking into account the characteristics exposed previously, we defined metrics for datawarehouses. These metrics can be applied at table, star or schema level. In this work we only show the metrics for datawarehouse schemas (table 1), that we have already validated theoretically in [9].

NFT(Sc). Defined as a number of fact tables of the schema.	**NDT(Sc).** Number of dimension tables of the schema.
NSDT(Sc). Number of shared dimension tables. Number of dimension tables shared for more than one star of the schema.	**NT(Sc).** Number of tables. Number of the fact tables plus the number of dimension tables of the schema.
NAFT(Sc). Number of attributes of fact tables of the schema. $NAFT(Sc) = \sum_{i=1}^{NFT} NA(FT_i)$ Where FTi is the fact table i of the schema Sc	**NADT(Sc).** Number of attributes of dimension tables of the schema. $NADT(Sc) = \sum_{i=1}^{NDT} NA(DT_i)$ Where DTi is the dimensional table i of the schema Sc
NASDT(Sc). Number of attributes of shared dimension tables of the schema. $NASDT(Sc) = \sum_{i=1}^{NSDT} NA(DTi)$ Where DTi is the dimensional table i of the schema Sc	**NA(Sc).** Number of attributes of the schema. NA(Sc) = NAFT(Sc) + NADT(Sc)
NFK(Sc). Number of foreign keys in all the fact tables of the schema. $NFK(Sc) = \sum_{i=1}^{NFT} NFK(FT_i)$ Where FTi is the fact table i of the schema Sc	**RSDT(Sc).** Ratio of Shared Dimension Tables. Quantity of dimension tables, which belong to more than one star. $RSDT(Sc) = \frac{NSDT(Sc)}{NDT(Sc)}$
RT(Sc). Ratio of tables. Quantity of dimension tables per fact table. $RT(Sc) = \frac{NDT(Sc)}{NFT(Sc)}$	**RScA(Sc).** Ratio of Schema Attributes. Number of attributes in dimension tables per attributes in fact tables. $RScA(Sc) = \frac{NADT(Sc)}{NAFT(Sc)}$
RFK(Sc). Ratio of foreign keys. Quantity of attributes that are foreign key. $RFK(Sc) = \frac{NFK(Sc)}{NA(Sc)}$	**RSDTA(Sc).** Ratio of Shared Dimension Tables Attributes. Number of attributes of the schema that are shared. $RSDTA(Sc) = \frac{NASDT(Sc)}{NA(Sc)}$

Table 1. Metrics for datawarehouse schemas

4. Example

Figure 2 shows an example of a Data Warehouse [1]. The values for the metrics are shown in table 2.

For example, we can see that the NA metric (Number of attributes) can be calculated as the sum of the number of attributes of each of the tables in the schema (NA = NA(LOT/SERIAL) + NA (COMPONENT) + NA(DEFECT) + NA(SHIPMENT) + NA(FACILITY) + NA(SUPLIER) + NA(TIME) + NA(COMP_RECEIPT_FACTS) + NA(DEFECT_FACTS) = 4 + 6 + 4 + 3 + 3 + 5 + 6 + 7 + 8 = 46)

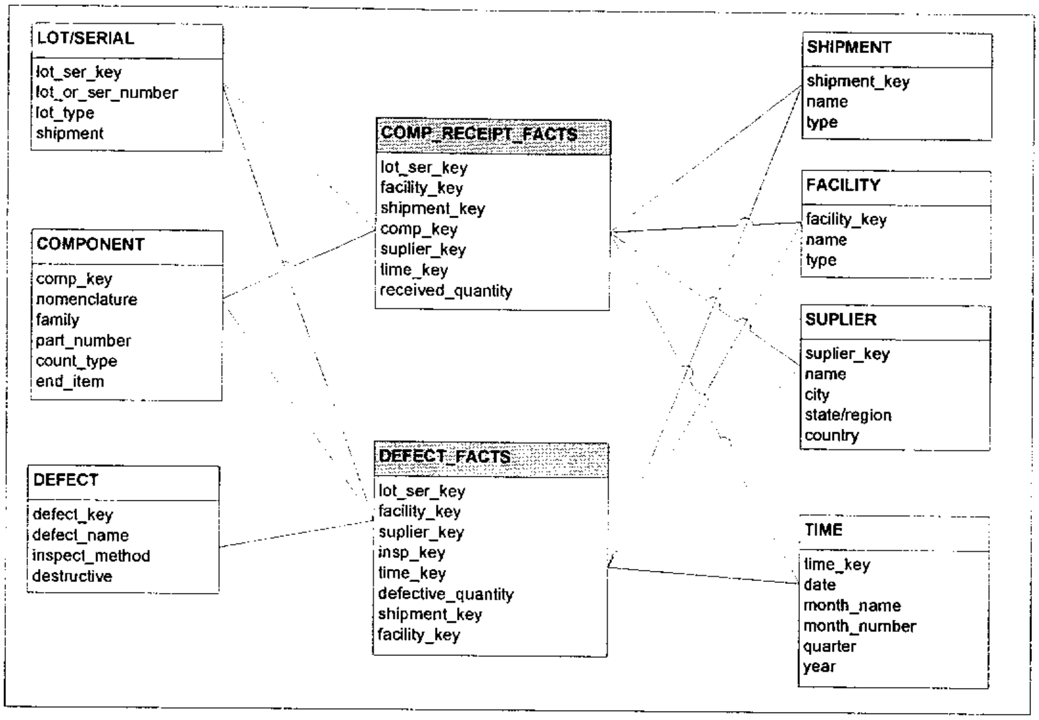


Figure 2. Example of a data warehouse star design [1]

Metric	Value
NA	46
NFK	13
NDT	7
NT	9
NADT	31
NAFT	15
RFK	13/46
NFT	2
NSDT	6
NASDT	27
RSDT	6/7
RT	7/2
RScA	31/15
RSDTA	27/46

Table 2. Values for the metrics

5. Empirical Validation of the Proposed Measures

In this section we describe an experiment we have carried out for empirically validating the proposed measures. We have followed some suggestions provided by [21], [18] and [6] about how to perform controlled experiments. For describing the experiment we use (with only minor changes) the format proposed by [21].

5.1. Definition

Using the GQM template [2] [3] for goal definition, the goal of the experiment is defined as follows:

Analyse	*DW structural complexity measures*
For the purpose of	*Evaluating*
With respect to	*the capability to be used as DW quality indicators*
From the point of view of	*Database designers*
In the context of	*Experts in database design*

5.2. Planning

- **Context selection.** The context of the experiment is a group of 12 experts in database design (Practitioners with an average of two years in database field). The experiment is specific since it is focused on DW structural complexity measures. The ability to generalise from this specific context is further elaborated below when discussing threats to the study validity. The experiment addresses a real problem, i.e. what indicators can be used to assess the quality of datawarehouses? To this end it investigates the correlation between DW structural complexity measures and the understandability of the datawarehouse (Thinking that an understandable DW schema has more quality).

- **Selection of subjects.** The subjects are chosen for convenience, i.e. the subjects are experts that have experience in database design, but they lack of knowledge in datawarehouse design.

- **Variables selection.** The independent variable is DW structural complexity. The dependent variable is the understandability of the datawarehouse schema.

- **Instrumentation**. The objects used in the experiment were eleven datawarehouse schemas. The independent variable was measured through the measures we proposed (see section 3). The dependent variable were measured by asking the subject to rank the understandability of each of the schemas (From 1 = too easy to 7=too complex).

- **Hypothesis formulation.** We wish to test the following hypotheses:
  - Null hypothesis, H_0: There is no significant correlation between the structural complexity measures and the understandability of the schemas.
  - Alternative hypothesis, H_1: There is a statistically significant correlation between structural complexity measures and the understandability of the schemas, that is also practically significant.

- **Experiment design.** We selected a within-subject design experiment, i.e. all the tests had to be solved by each of the subjects. The subjects were given the schemas in different order.

5.3. Operation

- **Preparation.** Subjects were given an intensive explanation session before the experiment took place. However, the subjects were not aware of the aspects we intended to study. Neither were they informed on the actual hypothesis stated.

  We prepared the material we handed to the subjects, consisting of eleven DW schemas. These diagrams are related to different universes of discourse that are general enough to be easily understood by each of the subjects. The structural complexity of each diagram is different, because as table 3 shows the values of the measures are different for each diagram.

Each diagram had a test enclosed that ask the subjects to rank the complexity of the schema.

Schema	NFT	NDT	NSDT	NT	NAFT	NADT	NASDT	NA	NFK	RSDT	RT	RScA	RFK	RSDTA
1	1	4	0	5	5	19	0	24	4	0,00	4,00	3,80	0,17	0,00
2	3	7	2	10	19	39	11	58	11	0,29	2,33	2,05	0,19	0,19
3	3	5	4	8	15	28	22	43	10	0,80	1,67	1,87	0,23	0,51
4	4	8	4	12	23	42	24	65	17	0,50	2,00	1,83	0,26	0,37
5	1	3	0	4	4	14	0	18	3	0,00	3,00	3,50	0,17	0,00
6	1	5	0	6	16	39	0	55	5	0,00	5,00	2,44	0,09	0,00
7	2	5	4	7	18	32	25	50	10	0,80	2,50	1,78	0,20	0,50
8	1	4	0	5	10	25	0	35	4	0,00	4,00	2,50	0,11	0,00
9	1	3	0	4	11	25	0	36	3	0,00	3,00	2,27	0,08	0,00
10	1	6	0	7	7	46	0	53	6	0,00	6,00	6,57	0,11	0,00
11	2	5	4	7	12	28	23	40	9	0,80	2,50	2,33	0,23	0,58

Table 3. Measure values for each DW schema

- **Execution.** The subjects were given all the materials described in the previous paragraph. We explained to them how to carry out the tests. We allowed one hour to do all the tests. Each subject had to work alone. In case of doubt, they could only consult the supervisor who organised the experiment.

- **Data Validation.** We collected all the tests controlling if they were complete and the responses were correct. All the test were correct.

5.4. Analysis and Interpretation

We used the data collected in order to test the hypotheses formulated in section 5.2. We made an statistical analysis of correlation using the Tau_b of Kendall statistic, with a level of significance $\alpha = 0.01$, which means the level of confidence is 99% (i.e. the probability that we reject H_0 when H_0 is false is at least 99%, which is statistically acceptable). Using the Tau_b statistic, each of the measures was correlated separately to the complexity of the schemas (see table 4).

Analysing the table 4, we can conclude that there is a high correlation (rejecting hypothesis H_0) between the complexity of the schemas and the measures NFK, NFT and NT because the correlation coefficient is greater than 0.5, which is a common threshold to evaluate correlation values. The metrics NAFT, NASDT, NDT, NSDT, RFK, RSDT and RSDTA seem to be less correlated to the complexity than the other measures (though the correlation value is still greater than 0.5). The metrics NA, NADT, RSCA and RT are not correlated with complexity.

			Complexity
Tau_b de Kendall	Correlation coeficient	NA	0,448
		NADT	0,365
		NAFT	0,518
		NASDT	0,520
		NDT	0,571
		NFK	0,664
		NFT	0,688
		NSDT	0,581
		NT	0,669
		RFK	0,520
		RSCA	-0,406
		RSDT	0,515
		RSDTA	0,507
		RT	-0,478

Table 4. Tau_b Kendall´s correlation coefficients between the measures and the complexity of the schemas

Even though the results obtained in this experiment are encouraging we cannot consider them as conclusive results. We are aware that it is necessary to replicate the experiment and to carry out new ones in order to confirm our results. Also it is necessary to apply these measures to data obtained from "real projects".

5.5. Validity evaluation

We will discuss the empirical study's various threats to validity and the way we attempted to alleviate them.

- **Threats to conclusion validity.** The conclusion validity defines the extent to which conclusions are statistically valid. One issue that could affect the statistical validity of this study is the size of the sample data (132 values, 11 diagrams and 12 subjects), which is perhaps not enough for both parametric and non-parametric statistic test [6]. We are aware of this, so we will consider the results of the experiment only as preliminary findings.

- **Threats to Construct Validity.** The construct validity is the degree to which the independent and the dependent variables are accurately measured by the measurement instruments used in the study. The dependent variable is subjective so, we must consider the experiment as a first approach. The construct validity of the measures used for the independent variables is guaranteed by the formal validation of the metrics we show in [9].

- **Threats to Internal Validity.** The internal validity is the degree to which conclusions can be drawn about the causal effect of independent variables on the dependent variables. The following issues have been dealt with:
  - DIFFERENCES AMONG SUBJECTS. Using a within-subjects design, error variance due to differences among subjects is reduced.
  - KNOWLEDGE OF THE UNIVERSE OF DISCOURSE. The DW schemas were from different universes of discourse but they are general enough to be easily

understood for each of the subjects. So that the knowledge of the domain does not affect the internal validity.

- LEARNING EFFECTS. The subjects were given the tests in different order, to cancel out learning effects. Subjects were required and controlled to answer in the order in which the tests appeared.
- FATIGUE EFFECTS. On average the experiment lasted for less than one hour, so fatigue was not very relevant. Also, the different order in the tests helped to cancel out these effects.
- PERSISTENCE EFFECTS. In order to avoid persistence effects, the experiment was run with subjects who had never done a similar experiment.
- SUBJECT MOTIVATION. We motivated subjects to participate in the experiment, explaining to them that the results of the experiment could benefit them as information systems practitioners.
- OTHER FACTORS. Subjects were told to not talk with each other.

– **Threats to External Validity.** The external validity is the degree to which the results of the research can be generalised to the population under study and to other research settings. The greater the external validity, the more the results of an empirical study can be generalised to actual software engineering practice. Two threats to validity have been identified which limit the ability to apply any such generalisation:

  - MATERIALS AND TASKS USED. In the experiment we tried to use DW schemas which can be representative of real cases, but more empirical studies using "real cases" from software companies must be done.
  - SUBJECTS. We are aware that more experiments with practitioners and professionals must be carried out in order to be able to generalise these results.

6. Conclusions and Future Work

If we really consider that information is "the" main organizational asset, one of the primary duties of IT professionals must be assuring its quality. Information quality can be decomposed in different types of "qualities": presentation quality, data warehouse management system quality, data quality, physical model quality and multidimensional model quality. Last one is our focus. Although some interesting guidelines have been proposed for designing "good" multidimensional models, more objective indicators are needed.

We are elaborating a set of valid metrics for measuring data warehouse quality, which can help designers in choosing the best option among more than one alternative design.

We have presented some metrics for measuring the data warehouse star design and we have explained the first experiment we have develop in order to validate the metrics, and we have obtained that some of them can be useful for our purpose. However, this is only a first approach and we must carry on more experiments to get a set of valid metrics.

Acknowledgements

This research is part of the CALDEA project (TIC 2000-0024-P4-02) supported by the Ministerio de Ciencia y Tecnología of Spain.

References

[1] ADAMSON, C. and VENERABLE, M.: 'Data Warehouse Design Solutions' (John Wiley and Sons, 1998)

[2] BASILI, V.R. and WEISS, D.: 'A methodology for Collecting Valid Software Engineering Data', *IEEE Transactions on Software Engineering*, 1984, SE-10. No. 6. pp. 728-738

[3] BASILI, V. and ROMBACH, H.: 'The TAME project: towards improvement-oriented software environments', *IEEE Transactions on Software Engineering*, 1988, **14**(6), pp. 728-738

[4] BASILI, V.R., SHULL, F. and LANUBILLE, F.: 'Building Knowledge through families of experiments', *IEEE Transactions on Software Engineering*., 1999, o. 4. pp. 456-473.

[5] BOMAN, M., BUBENKO, J., JOHANNESSON, P. and WANGLER, B.: 'Conceptual Modelling' (Prentice Hall, 1997)

[6] BRIAND, L., EL EMAM, K. and MORASCA, S.:. 'Theoretical and empirical validation of software product measures'. Technical report ISERN-95-03, International Software Engineering Research Network, 1995

[7] BRIAND, L.C., MORASCA, S. and BASILI, V.: 'Property-based software engineering measurement', *IEEE Transactions on Software Engineering*., 1996, **22**(1) pp. 68-85

[8] CALERO, C., PIATTINI, M. and GENERO, M.: 'Developing quality complex database systems: practices, techniques and technologies', (Becker, Idea Group Publishing, 2001)

[9] CALERO, C., PIATTINI, M., PASCUAL, C., SERRANO, M.: 'Towards Datawarehouse Quality Metrics'. International Workshop on DESIGN AND MANAGEMENT OF DATA WAREHOUSES, DMDW'01 in conjuntion with the 13th Conference on Advance Information Systems Engineering, CAISE'01, June, 2001

[10] FENTON, N. and PFLEEGER, S.: 'Software Metrics: A Rigorous Aproach' (Chapman & Hall, London, 1997, 2nd EDN.)

[11] GARDNER, S.R.: 'Building the data warehouse', *Communications of the ACM,* September, 1998, Vol. 41, Nr.9. pp. 52-60

[12] HAMMERGREN, T.: 'Data Warehousing Building the Corporate Knowledge Base' (International Thomson Computer Press, Milford, 1996)

[13] INMON, W. H.: 'Building the Data Warehouse' (John Wiley and Sons, 1997, 2nd edn.

[14] JARKE, M., LENZERINI, M., VASSILOU, Y. and VASSILIADIS, P.: 'Fundamentals of Data Warehouses' (Ed. Springer, 2000)

[15] KELLY, S.: 'Data Warehousing in Action' (John Wiley & Sons, 1997)

[16] KIMBALL, R., REEVES, L., ROSS, M. and THORNTHWAITE, W.: 'The Data Warehouse Lifecycle Toolkit' (John Wiley and Sons, 1998)

[17] LINDLAND, O., SINDRE, G. and SOLVBERG, A.: 'Understanding Quality in Conceptual Modelling', *IEEE Software,* March 1994, Vol. 11 N° 2, pp. 42-49

[18] PERRY D., PORTER A. and VOTTA L.: 'Empirical Studies of Software Engineering: A Roadmap. Future of Software Engineering' (Ed:Anthony Finkelstein, ACM, pp. 345-355, 2000)

[19] WEYUKER, E.J.: 'Evaluating software complexity measures', *IEEE Transactions on Software Engineering*, 1988, **14**(9) pp. 1357-1365

[20] WHITMIRE, S.A.: 'Object Oriented Design Measurement' (Ed. Wiley, 1997)

[21] WHOLIN, C., RUNESON, P., HÖST, M., OHLSON, M., REGNELL, B. and WESSLÉN, A.: 'Experimentation in Software Engineering: An Introduction' (Kluwer Academic Publishers, 2000)

[22] ZUSE, H.: 'A Framework of Software Measurement' (Walter de Gruyter, 1998)