# AN INTEGRATIVE FRAMEWORK TO ASSESS AND IMPROVE INFORMATION QUALITY MANAGEMENT IN ORGANIZATIONS

Ismael Caballero, Jesús Rodríguez, Mario Piattini
*ALARCOS Research Group.*

Keywords: Data Quality, Information Quality, Data and Information Quality Management

Abstract: Information quality has become a decisive factor in organizations since it is the basis of strategic decisions. Many researching lines over the last decade have looked at specific data and information quality issues from different standpoints. Taking care about data and information quality goes beyond the definition of data quality dimensions, and today, there is still a lack of an integrative framework, which can guide organizations in the assessment and improvement of data and information quality in a coordinated and global way. This paper tries to fulfil this gap by proposing a framework which uses the Information Management Process (IMP) concept. It consists of two main components: an Information Quality Management Model structured in Maturity Levels (CALDEA) and an Assessment and Improvement Methodology (EVAMECAL). The methodology allows the assessment of an IMP in terms of maturity levels given by CALDEA, which is used as a guidance for improvements.

## 1 INTRODUCTION

It is a widely known fact that dealing with information problems is not a trivial issue, nor free. Many resources are required because quality assurance is a complex process, in which the difference between costs and required quality is closely linked to the context of the application and organization requirements (Bringel et al., 2004). Nowadays most of the pragmatic and theorical efforts achieved on information quality researches are focused on solving specific and concrete problems regarding to technical or managerial information quality issues. These efforts often lack of an strategic perspective that does not allow organization to optimize the effectiveness of their information quality initiatives in an organizational scope. Many organizations, even after having identified their information quality problems, do not have the correct techniques, tools and practices to implement some of the proposed solutions through researching lines. Information quality issues are not usually understood as a global problem of the entire organization, but a punctual and an isolated one. It might be a matter of a quality management team, encouraged by organization heads, who must implement several quality management concepts like information quality policy, information strategy, information quality planning, information quality control and information quality assurance through the organization; implying all workers by commitments and trying to coordinate efforts and resources in order to control and improve information quality issues with a strategic perspective. Unfortunately, there is not still an integrative framework that guides organizations to achieve information quality goals through management by implementing the concepts mentioned above

Trying to fulfil this lack, we are going to propose an integrative framework considering information as a product – which allows to take an engineering point of view for information quality-, and taking into account the Software Process definition given by Fuggeta (2000) - which allows to identify who, when and how is using whatever resources to view both Information Management and Information Quality Management activities as an Information Management Process (IMP), in order to model what happens in organization and how information quality might be managed. Information quality is going to be managed by assessing and improving a concrete

Caballero I., Rodríguez J. and Piattini M. (2005).
AN INTEGRATIVE FRAMEWORK TO ASSESS AND IMPROVE INFORMATION QUALITY MANAGEMENT IN ORGANIZATIONS.
In *Proceedings of the Seventh International Conference on Enterprise Information Systems - DISI*, pages 403-406

IMP. It is true that there are several frameworks for assessing and improving software processes such as CMM, CMMI, ISO 9001, BootStrap, and SPICE; but none of them have focused on information quality nor even taken it into account.

Our proposal defines two main components: An Information Quality Management Model, (CALDEA) which serves as a reference model when using the second component, the Assessment and Improvement Methodology (EVAMECAL).

The remainder of this paper is structured as follows: In Section 2, The Information Quality Management Process is being shown. The main steps of the Assessment and Improvement Methodology and regarding issues are being drawn in Section 3. Finally, In Section 4, some conclusions are going to be highlighted.

## 2 CALDEA: AN INFORMATION QUALITY MANAGEMENT MODEL

CALDEA takes the maturity-staged levels from CMMI and defines five information quality maturity levels for an IMP as well as CMMI: Initial, Definition, Integration, Quantitative Management and Optimizing. Each level stands for specific information quality management goals. For each maturity level, several Key Process Areas (KPA) are proposed. These KPAs are not only focused on technical but also managerial issues, providing the basis for information quality measurement and management and linking both aspects. For each KPA, some activities, tools, techniques, standards, practices, and metrics as required, are proposed, but not imposed, in order to make the model as universal and general as possible. This structure of maturity levels allows organizations to take an strategic perspective for the efforts achieved. The maturity levels and corresponding KPAs are:

- **Initial Level:** An IMP is said to be at Initial Level when no efforts are made in order to achieve any information quality goals.

- **Definition Level**: An IMP is said to be at Definition Level or Defined when it has been defined and planned. This implies identifying all its components and their relationship with the requirements. In order To achieve this goal, the following KPAs need to be satisfied:

  - **(IQATM) Information Quality Assurance Team Management.** The aim of this KPA is to form a team composed by people who have direct responsibility on information and on its integrity. This team will encourage the entire organization to take on commitments regarding information quality policies (Ballou and Tayi, 1999) and to make corresponding efforts in order to support the activities of this maturity model.
  - **(IPM) IMP Project Management.** This is a management KPA aimed at developing a plan for IMP in order to coordinate both managerial and technical efforts and to elaborate all the documentation related
  - **(URM) User Requirements Management.** User Requirements Specification (URS) must be collected and documented. Three kinds of requirements might be identified: those related to final information product (URS-IP), those related to IMP – which must be gathered in the User Requirement Specification for IMP document (URS-IMP) document - and those related to Information Quality –which must be gathered in the Information Quality User Requirements Specification (URS-IQ).
  - **(DSTM) Data Sources and Data Targets Management.** Both data sources and targets must be identified and documented, in order to avoid problems such us uncontrolled data redundancy or problems with data format interchange.
  - **(ADMPM) Database or Data Warehouse Acquisition, Development or Maintenance Project Management**. This should support both URS-IQ and URS-IMP.
  - **(IQM) Information Quality Management in IMP Components**. It is necessary to identify from the URS-IQ the dimensions of quality of information that must be controlled (Huang *et al.*, 1999), as well as the metrics adapted for each one of those dimensions (Kahn *et al.,* 2002).

- **Integration Level** An IMP is said to be at Integration Level or Integrated when after being having been Defined (Definition level has been achieved), many efforts are made in order to assure that the IMP is in compliance with organizational information quality requirements and standards. This implies standardizing different information quality learned lessons in order to avoid previous mistakes and to improve future work. The following KPAs must be satisfied:
  - **(VV) Information Products and IMP Components Validation and Verification**. Both information products (obtained as a result of data transformation process) and

IMP components must be verified and validated to correct defects and/or discord with the USR-IMP, USR-IQ and the organizational information quality policies.

- **(RM) Risk and Poor Information Quality Impact Management.**
- **(IQSM) Information Quality Standardization Management**. All lessons learned through specific experiences should be properly gathered, documented .
- **(OIQPM) Organizational Information Quality Policies Management**. The way by which all the efforts previously mentioned can be implemented.

**- Quantitative Management** The main information quality goal of this level is to obtain a quantitative compliance that IMP performance over a reasonable time period, remains as consistent as required in terms of variation and stability through a reliable set of measurements of information quality characteristics of IMP. This level is composed by the following KPA:

- **(MM) IMP Measurement Management.** Since metrics about IMP components have been drawn up at definition level, the aim of this KPA is to define when and how to make the measurements, how to represent the results and to whom.
- **(AMP) IMP Measurement Plan Automation Management**. This KPA aims to study all the issues related to the automation of these measurement procedures.

**- Optimizing Level**. An IMP is said to be at Optimizing Level if when being quantitatively managed the obtained measurements are used to develop a continuous improvement by eliminating defects or by proposing and implementing several improvements. The following two KPAs must be satisfied:

- **(CADPM) Causal Analysis for Defect Prevention Management**. From the study of the measurement results, some typical quality techniques and tools like Statistical Control Process (SPC) or Ishikawa's diagrams can be applied to detect defects of information quality and identify their root causes.
- **(IODM) Innovation and Organizational Development Management.** Similarly to the previous KPA, the results can be also used here to improve the IMP, in terms of performance, planned time or budget. This is the basis of the idea of continuous improvement.

# 3 EVAMECAL: AN ASSESSMENT AND IMPROVEMENT METHODOLOGY

**EVAMECAL** is based on Deming's continuous improvement cycle PDCA (*Plan-Do-Check-Act*). The main aim of EVAMECAL is to assess and to improve a specific IMP of a given organization. Assessments and improvements results are made in terms of maturity levels given by CALDEA.

Taking as a reference the evaluation model from ISO/IEC 15504 we have drawn a set of possible states for maturity levels: {"Consolidated" / "No Consolidated"}. We have also set states for KPA´s, activities and Components.

Additionally, we have drawn a set of rules to determine the state of each element, which are based on the concept of Information Quality Value (IQV), which can furthermore be computed as a weighted average of the Criticality Degree. The rules also establish the value ranges of each state for each element.

## 3.1 Steps and Activities of EVAMECAL

An improvement program starts with the recognition of the needs and goals of the organization in order to determine the improvement objectives. The improvement program should also reflect the present situation of the IMP and the main efforts, organized in an Improvement Plan for the IMP (IP-IMP), to be made in order to reach the objectives. After the execution of the IP-IMP, it is necessary to check the correcting actions executed and to develop a report about the experience of the plan so the know ledges can be useful to avoid future mistakes . This phases or steps can be grouped into four blocks of a PDCA cycle.

**(EMC-P) EVAMECAL-"PLAN"**

The planning phase consists of the following sub phases:

- **Definition of actual state of IMPs**. This is the assessment step. The main aim is to determine at which maturity level an IMP is. The scope of this activity can be defined as the measurement of the IQVs which is made by using the defined questionnaires and some other tools.
- **Definition of Improvement Goals.** Having into account the obtained results at the assessment step, next step is to define the scope of the improvements. This implies to establish a set of improvement goals in terms

of the states of each element previously described. In order to realize this activity the actual state of the IMP should be compared with the model proposed by CALDEA, so that the activities, --which need to be executed to improve and to achieve an objetive--, can be identified.

**(EMC-D) EVAMECAL-"DO"**

**Causal Analysis for Defect Prevention.** In order to reach the proposed improvement goals, it is necessary to determine the source of detected defects trying to remove the gap between actual and desired state. This activity aims to the design of tests which allow to detect the defects.

- **Definition and execution of an Improvement Plan for the IMP.** This is the improvement step. Oncedefect sources have been identified, an Improvement Plan for the IMP (IP-IMP) containing corrective actions must be defined. It is also important to manage the associated risks, the total cost of the improvement project and the benefits so that the viability of the Plan can be studied. If the IP-IMP is viable, it is executed.

**(EMC-C) EVAMECAL-"CHECK"**

- **Checking for effectiveness of the Improvement Plan.** In order to check if Improvement Plan has worked properly, a new assessment like in step 1 is required. If improvements goals have been reached, then go to step 6. Otherwise, go back to step 3.

**(EMC-A) EVAMECAL-"ACT"**

- **Get conclusions and standardize the learned lessons.** This implies the study of the gap between the initial prediction of resources and benefits, and the result of the taken actions.

# 4 CONCLUSSIONS AND FUTURE WORK

In this paper, the concept of IMP and a framework to optimize information quality in organizations have been presented. It consists on two elements: an Information Quality Management Model (CALDEA) and an Assessment and Improvement Methodology (EVAMECAL). The way to use this framework may be stated as follows: first, identify the IMPs of the organization and choose the most critical ones; second, apply EVAMECAL for assessing and improving the chosen IMPs. Assessment and improvement sequences are going to be made having CALDEA as reference.

These components satisfy the conditions proposed by Eppler and Wittig (2000) for a good information quality framework: CALDEA provides a systematic and concise set of criteria for information quality according to which information can be evaluated. EVAMECAL provides a schema to analyze and solve information quality problems. CALDEA is by itself a conceptual map that can be used to structure a variety of approaches, theories and information quality related phenomena since KPA does not propose a closed set of tools, techniques and methodologies.

# ACKNOWLEDGEMENTS

This research is part of both CALIPO- supported by Dirección General de Investigación of the Ministerio de Ciencia y Tecnología (TIC2003-07804-C05-03)- and MESSENGER project - supported by Consejería de Ciencia y Tecnología de la Junta de Comunidades de Castilla-La Mancha (PCC-03-003-1).

# REFERENCES

Ballou, D. and Tayi, G.K. "Enhancing data quality in Data Warehouse Environments". *Communications of the ACM*, January 1999/ Vol 42, No I.

Bringel, H., Caetano, A., Tribolet, J. Business Process Modeling Towards Data Quality Assurance. *Proceeding of ICEIS'2004.* pp. 565-568. Porto, Portugal. 2004

Eppler, M.J. and Wittig, D. "Conceptualizing Information Quality: A review of Information Quality Frameworks from the last ten years." *Proceedings of the 2000 Conference on Information Quality*. Pp 83-96

Fuggeta, A. "Software Process: A roadmap. The future of Software Engineering", ed. A. Finkelstein *ACM, Press*, 2000, pp.27-34.

Grimmer, U., and Hinrichs, H. "A methodological approach to data quality management supported by data mining". *Proceedings of the Sixth International conference on Information Quality*. 2001 Pp 217-232

Huang, K.T., Lee, Y., Wang, R. *Quality Information and Knowledge*. Prentice Hall, Upper Saddle River, 1999

Kahn, B., Strong, D., Wang, R. "Information Quality Benchmarks: Product and Service Performance". *Communications of the ACM* April 2002/Vol. 45, No 4

Xu, H., Would organization size matter for data quality. *Proceedings of the Eighth Conference on Information Quality ICIQ'2003*. Pp. 365-379. 2003.