Towards a Data Quality Framework for Decision Support in a Multidimensional Context

Daniel Poeppelmann, University of Osnabrueck, Germany
Christian Schultewolter, University of Osnabrueck, Germany

ABSTRACT

Data quality is defined as fitness for use; meaning it is strongly influenced by contextual factors. None of the proposed methods for data quality assessment considers the management level and decision process phase as important contextual factors. Each combination of different management levels and decision process phases can be described as the overall context in which data quality is to be assessed and leads to a certain information demand. The information demand can in turn be described by a set of features whose values depend on the overall context. A context-sensitive assessment of data quality should therefore be based on data quality dimensions related with certain features characterizing the information demand and their values. This paper proposes a framework for data quality assessment based on the contextual factors management level and decision process phase which allows a decision maker to relate data quality dimensions to values of features of the information demand. Based on this individual assignment of dimensions, the user can access situation specific information regarding data quality.

Keywords: Business Intelligence, Case Study, Context-Sensitivity, Data Quality, Decision Process, Decision Support, Management Level

MOTIVATION

Data quality directly influences decisions on all management levels of enterprises (Redman, 1998). Especially on tactical and strategic levels decision processes are heavily dependent on the quality of underlying data (Price & Shanks, 2005). Thus, data quality plays an important role in a wide variety of information systems applications. From a data consumer point of view it is defined as fitness for use (Wang & Strong, 1996). This definition emphasizes the role of the data consumer who decides whether data is of an appropriate quality for usage in a given context.

Literature provides evidence that supplying decision makers with information regarding the quality of available data positively influences decisions (Chengular-Smith, Ballou, & Pazer, 1999; Even, Shankaranarayanan, & Watts, 2006; Shankaranarayanan, Ziad, & Wang, 2003; Shankaranarayanan & Cai, 2006). According to the definition of Wang and Strong (1996) the context of the decision maker has to be taken...
into account when measuring data quality and generating information about data quality. Current literature does not consider the context in an adequate manner. Especially the division of a decision processes into four phases (Simon, 1977) and the distinction of three management levels strategic planning, management control and operational control (Anthony, 1965) on which decisions have to be made are not reflected by existing approaches. The combination of the context dimension A: decision process phase and context dimension B: management level determines the overall context in which a decision maker finds himself (Figure 1). As each of the decision process phases and management levels come along with different information needs (Gorry & Scott-Morton, 1971) consequently each overall context requires special information, too. In conjunction with the definition of Wang and Strong (1996) data quality possibly defers in relation to a given combination of both context dimensions. The decision maker as data consumer has to decide on the fitness for use of the underlying data with regard to the overall context he is faced with and the resulting information need.

This paper proposes a framework for the selection of adequate dimensions of data quality for different overall contexts a decision maker may be faced with. It is intended to be used to support decision makers with context-sensitive information regarding data quality. Thus we serve the claim for differentiated needs for data quality information in specific decision-situations with regard to the context dimensions management level and decision process phase. This will serve the overall goal to improve decision outcomes.

Our research methodology is based on the approaches of Rossi and Sein (2003) and Hevner, March, Park, and Ram (2004). Currently there is no approach for context-sensitive data quality assessment regarding decision making phase and management level (need). Thus we design a framework (artifact) and describe the application with a fictitious example. We claim that decision processes on each management level and in each decision process phase can be enhanced when applying our framework (hypothesis 1) and that different combinations of management level and decision making phase come along with different requirements regarding data quality (hypothesis 2).

The next section explains basic characteristics of decisions and decision processes. It gives an overview of the differing information needs in each decision process phase and on each management level. We then examine state-of-the art of data quality research and especially identifies appropriate data quality dimensions and metrics. Composing the insights of these sections namely differing information needs and possibilities of measuring data quality as means to satisfy these needs, we propose a framework for data quality assessment in given overall contexts. Afterwards we provide an example application of the framework.

**Dimensions of Decision Making Contexts and Information Demands**

**Context-Dimension A: Phases of Decision Processes**

Decision making can be described as a process of choosing one out of a given set of alternative ways of action. To define a problem as a decision-problem there has to be a set of alternatives that can be ranked in terms of goal attainment (Turban, Sharda, & Delen, 2010). Otherwise the problem would be a choice-problem rather than a decision-problem as one could choose any of the given alternatives without any impact on goal attainment. For the evaluation of the alternatives the decision-maker depends on information. Thus the quality of available information directly influences the quality of the decision itself.

According to Simon (1977) a decision is a process composed of the four phases intelligence activity, design activity, choice activity, review activity whereat in any of the latter three phases there may be a return to an earlier phase. The intelligence activity identifies a decision-problem and gives a detailed description by
Business Intelligence Practices: Adding Evidence from Organizations in the Nordic Countries
www.igi-global.com/article/business-intelligence-practices/78273?camid=4v1a