Data Quality in eMaintenance: A Call for Research

Mustafa Al-Jumaili[#1], Ville Rauhala [*2], Phillip Tretten[#3], Ramin Karim[#4]

[#] *Division of operation and maintenance engineering,*
Luleå University of Technology, Luleå, Sweden
[1]mustafa.aljumaili@ltu.se
[3]phillip.tretten@ltu.se
[4]ramin.karim@ltu.se

[*] *Kemi-Tornio University of Applied Science, Technology*
Kemi, Finland
[2]ville.rauhala@tokem.fi

***Abstract*— Effective and efficient maintenance requires a proper information logistics, which can be delivered through eMaintenance solutions. Development of eMaintenance solutions faces extensive challenges. One of these challenges is how to ensure the quality of data used in different eMaintenance solutions. Data Quality (DQ) concerns all phases of the maintenance process. The purpose of this paper is to answer the research question: how should DQ be considered and managed when developing eMaintenance solutions. To deal with such challenges a case study was conducted at a mining company. Empirical data has been collected through interviews, observations, archival records and workshops. The data analysis has been based on an empirical framework that supports the identification of required information services. Conditions that support the DQ and the information logistics, along with that, support the maintenance process have been presented. These aspects have also been related to the phases of a generic maintenance process.**

***Keywords*— eMaintenance, maintenance process, information logistics, Data Quality, Information and Communication Technology, Information Quality, Information System, Computerized Maintenance Management Systems.**

I. Introduction

eMaintenance can be defined as a maintenance strategy where tasks are managed electronically using real time data obtained through digital technologies (i.e. mobile devices, remote sensing, condition monitoring, knowledge engineering, telecommunications and internet technologies) [1]. eMaintenance is also considered as a maintenance plan to meet the productivity through condition monitoring, proactive maintenance and remote maintenance through real-time information support for decision making. eMaintenance thus can be termed as a maintenance support for the e-operation through remote diagnostics and asset management, simulation for optimisation and decision making under an e-business scenario for an organisation. As a provider and consumer relationship scenario, the term eMaintenance is often related to providing maintenance services remotely by enhanced use of ICT [2]. However, eMaintenance solutions aim to facilitate maintenance decision-making by providing an effectiveness and efficiency in maintenance information logistics.

When dealing with maintenance information logistics content management (i.e. 'what to deliver') is highly important, since it addresses issues related to the content including data and information [2]. However, in a decision-making process the quality of decision is strongly linked to the quality of the underlying data used during the data analysis, [3]. Hence, aspects related to DQ need to be emphasised in solutions aimed to support the decision-making process. Therefore, in development of eMaintenance solutions, DQ is essential and it needs to be considered the solution's whole lifecycle.

DQ in maintenance forms the integral part of information logistics, which is the backbone of a maintenance system derived from a clear maintenance strategy that, in turn, should be derived from and linked to the corporate strategy. This paper presents a broader perspective and contributes to DQ issues in all phases of the maintenance process. The purpose of this paper is to answer the research question: how should DT be considered and managed when developing eMaintenance solutions.

II. Theoretical Framework

The management of maintenance consists of activities as: developing and updating the maintenance policy, providing finances for maintenance and coordinating and supervision of maintenance (see Figure 1). The elements of maintenance support planning are: maintenance support definition, maintenance task identification, maintenance task analysis and maintenance support resources. Maintenance preparation concerns the planning for specific maintenance tasks, which includes planning of maintenance tasks, scheduling activities and assigning and obtaining resources. The maintenance execution phase includes the actual performance of maintenance, recording results and special safety and environmental procedures. Maintenance assessment includes measurement of maintenance performance, analysis of results and assessment of actions to be taken. Finally, maintenance improvement is achieved by improving the maintenance concept, the resources, the procedures and the equipment.

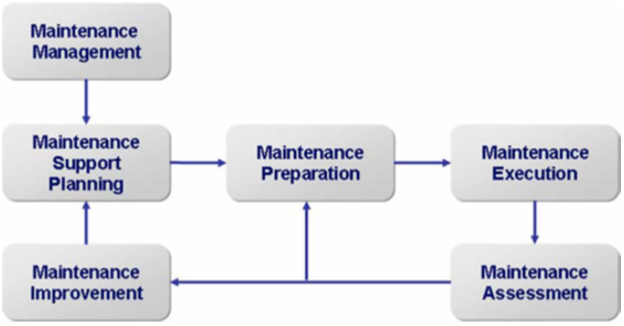

FIGURE 1. PHASES OF AN OVERALL MAINTENANCE PROCESS [4]

A. eMaintenance

The eMaintenance concept was introduced in the early 2000 and is today a common term in maintenance-related literature [5]. However, there are different views of what eMaintenance actually is. Some consider eMaintenance as a maintenance strategy, where tasks are managed electronically by the use of real-time item data obtained through digital technologies [1]. From a general perspective eMaintenance, concerns the use of new ICT in the maintenance area [6]. In this paper we adopt a broad view on eMaintenance in line with [5], as maintenance processes managed and performed via computing. This definition thus includes maintenance activities in all phases of the maintenance process and it includes a variety of ICT solutions ranging from computerised maintenance systems to sensor technologies. eMaintenance has gained increased attention in research with emphasis on a variety of issues. According to [7] four general issues addressed in eMaintenance literature: 1) standards, 2) platform development, 3) process formalization, and 4) system development and implementation. Besides these issues the literature also addresses eMaintenance from an information logistics perspective [1] [8] [9].

B. Data Quality

Quality is the degree to which a set of inherent characteristics fulfils requirements [10]. Hence, Data Quality (DQ) can be described as the degree of a set of permanent characteristics in the data or dataset that fulfils a certain data consumer's requirements or stated expectations. Furthermore, information can be described as the meaningful data [10]. Simultaneously, Information Quality (IQ) can be described as the degree of a set of permanent characteristics in the information that fulfils a certain consumer's requirements or stated expectations

For most contemporary organizations quality management is well-known and integrated into the everyday work for improving the quality of their product or service. However, when it comes to quality of the information that is produced and consumed in the organization the quality focus is not that extensive. Wang argues that the field of product quality has an extensive body of literature on Total Quality Management while information quality not has such an extensive body of literature [11]. High quality information is according to Wang dependent on the quality of the raw data and the process throughout which is processed [11].

However, maintaining data at a high quality level involves significant costs. These costs are associated with efforts to detect and correct defects, set governance policies, redesign processes, and invest in monitoring tools. From an economic perspective; one would try to reach a certain quality level at a minimum possible cost [12].

There are a number of theoretical frameworks for understanding DQ [13] [14] [15]. Wand and Wang have defined a set of DQ dimensions [15]. These are four intrinsic (system-oriented) DQ dimensions to specify whether data is complete, unambiguous, meaningful, and correct. While Price and Shanks have suggested a semiotic framework that classifies data into levels: form, meaning and application according to the use of data and quality needed [14]. Different techniques to improve the DQ have also been suggested, such as data profiling, data standardization, linking and data cleaning. DQ is thus of importance both when designing an eMaintenance solution and during its operational phase.

Information Quality (IQ) also used as a description to the DQ where "information" refers to both data and information. Some of literature has used both expressions to represent DQ making it difficult to distinguish between the two expressions. In the other hand, some authors tried to distinguish between DQ and IQ. According to Rosanne and Graeme, DQ is product-based perspective and hence the quality is defined in terms of the degree to which the IS data meets basic requirements specifications, while IQ is a service-based perspective of quality which focuses on the response of the consumer to their interactions with the IS [16]. Thus, IQ and DQ are intrinsically related and in this paper we will focus on DQ as an important enabler of IQ.

In this paper we will adopt the DQ dimensions suggested by Wand and Wang [15] to classify the DQ problems found in our case study.

III. DATA QUALITY IN EMAINTENANCE

The quality of the data in manufacturing has not been thoroughly discussed in prior research [17]. Moreover, only a limited number of papers focus on DQ and areas of maintenance.

Data quality issues are also considered in different industrial applications. Problems related to data quality are often handled with different software tools supporting to produce good quality data in information systems and applications such as OPC and SCADA.

IV. THE CASE STUDY

The case study was conducted at a Scandinavian mining company with approximately 160 employees. The plant consists of the actual mine and a concentrator plant. This case study focuses on the maintenance process for the concentrating plant, since maintenance activities in the mine are mainly conducted by subcontractors. The concentrator plant is led by a manager, who has two section managers, one responsible for production and one for plant maintenance and automation. The maintenance and automation function is divided into a mechanical and an electrical team.

In this case study a first step was to get a view of the current state of the maintenance process in the concentrating plant. For this, documentation provided by the mining company was studied to understand their organizational structure and the operational process more precisely. After that, we conducted ten interviews with people working at the concentrating plant. We sought to interview people from different positions and organization levels to get as many perspectives on the maintenance process as possible. The interviews were conducted in two rounds, in the first round six interviews were conducted with maintenance and production personnel. In the second round, four interviews were conducted with maintenance technicians to understand the information flows and IT support for maintenance execution. Table 3 shows titles and tasks of the interviewed personnel from maintenance and production organisations.

TABLE 1.
TITLES AND TASKS OF INTERVIEWED PERSONNEL

Title	Task
Section Manager	Automation and plant maintenance
Foreman	Mechanical maintenance
Mechanic	Mechanical maintenance
Foreman	Electrical/automation maintenance
Mechanic	Electrical/automation maintenance
Engineer	Development of mechanical maintenance, Condition monitoring
Technician	Preventive maintenance, KaTTi upkeep
Mechanic	Preventive maintenance
Operator	Production
Designer	Automation systems

The interviews were semi-structured with open-ended questions and lasted about one hour each. The interview structure was based on the sub phases of the maintenance process (Figure 1). These phases guided the semi-structured interviews, but we were open to pursuing follow-up questions and issues raised by the respondents. All interviews were tape recorded and followed up during the data analysis. After the interviews, the problems related to information flows in maintenance were compiled and structured according to their relevance to DQ and impact on the different phases in the maintenance process.

V. RESULTS

The ICT solutions used in the maintenance process at the studied company were, with few exceptions, a number of “stand alone” maintenance management systems with poor data exchange. Most of the data had to be manually transferred between systems. In addition to these systems, a large proportion of the maintenance information flows were supported by phone. The aim of the interviews was to get an understanding of how the different maintenance phases were supported by ICT and the information flows throughout the phases. In the study a number of problems were identified related to the use of ICT in the maintenance work.

A. Identified Issues

In the case study we have identified a number of problems related to the use of ICT in the maintenance phases. We first present all identified problems and then identify which ones that are related to DQ issues.

1) *Stand-alone IT-systems:* The CMMS does not contain all information required for the operative and strategic decision making in the different phases of the maintenance process. Also there are different sections in maintenance information systems, which are not connected with each other. This causes problems to integrate data to produce valuable information, for example maintenance management. Also information search is difficult in this case, for example finding spare parts and documents. Data in production, material management, warehouse and purchasing related IT-systems were not connected to data in the maintenance information system. The data was not comparable and it was difficult to combine, complicating maintenance related operative and strategic decision-making.

2) *Data Multiplicity:* This problem concerns the duplication of information and its existence in different systems or parts of a system. Certain information regarding technical documentation and drawings are stored both in the maintenance system and in other media (i.e. computer hard-drives, CD-ROMs, paper). However, in many cases there was no technical information, drawings or work instructions in the maintenance system. The stored information was not regularly updated or maintained and, thus, making it difficult to know if it can be trusted. Sometimes, the information sources contained conflicting information, which made it difficult to trust.

3) *Manual input and transfer of data:* This problem relates to the practice of feeding the systems with information. The practice of manually entering data has led to a wide variation of data quality, e.g. some of the maintenance personnel enter very detailed and descriptive failure reports, while others can omit compulsory information. Thus, the maintenance foreman often had to consult the operational staff and do a physical inspection to get necessary information.

4) *Usability of systems*: This problem points to the interface and functionality of the systems used in the maintenance processes. Several shared the opinion that the usability of the maintenance systems could be improved. Since it was difficult to find relevant information and user guidance was lacking, as well.

5) *IT-solutions are not available on the work site:* Maintenance related IT-solutions are made to improve different parts of the maintenance process. However IT-solutions were not commonly available at the jobsite. The most common situation was that maintenance and operative staff used the IT-solutions through the office PC. Although the most significant part of the data was collected from the jobsite

while maintenance was execution (see figure 1). Along with that they wanted maintenance related information, via IT-solutions, on the jobsite.

6) Performance indicators missing or difficult to obtain: To continuously follow up and improve the maintenance organization, it is important to have indicators of its status and efficiency. Maintenance assessment is currently based on information from SAP, where material costs, salaries for the maintenance staff and costs for subcontractors are followed up. However, material costs can only be followed on a sub process level, not on equipment level. Moreover, salaries are only followed at the plant level. There was no information available on the efficiency of the maintenance organization (due to invalid data of work time execution and status of work orders). Thus, it was difficult to make assessments and improvements of the organization based on real performance.

7) *No instructions or guiding:* One problem related to DQ is the lack of instructions and education for operative and maintenance staff concerning data collection. There was no plan or common practice, for what information, by whom and in what form should be collected in different phases of the maintenance process. There were several different methods to collect data and, therefore, information in the database was not compatible. Making it difficult to do systematic data-analysis and to produce key performance indicators.

8) No data upkeep: One normal problem concerning DQ in information systems is data upkeep. It has been discovered that without systematic upkeep the quality of data in information systems will decrease. To produce reliable information from large amounts of data, data has to be correct and in correct form. This requires regular checks and adjustment of the data. Along with that, regular data upkeep makes it is possible to identify and correct problems in the data collection phase. In this case study there was no person responsible for data upkeep.

9) *Problems with master data*: Usually when problems with DQ are discussed, it concerns transaction data. Transactions data means data collected from different parts of maintenance process during normal work tasks. However one problem concerns the quality of master data. Master data means data, which is, mainly created during the implementation phase of the information system, i.e. equipment hierarchy, position numbers, device numbers and spare parts. Gaps or mistakes in master data, significantly affects the quality of the collected transaction data. Good quality master data, created in implementation phase, is a prerequisite to good quality transaction data and it also needs frequently upkeep to maintain good quality.

10) *Data not used in decision making*: In this case study CMMS was mainly used for information transfer and communication between people of different positions and organizations, e.g. sending and receiving work orders or accessing information from the system. The data in CMMS was not collected for operative and strategic decision making. For this reason people were not motivated to collect good quality data in its correct form because they were not aware of what purpose the data was to be used for.

11) *Knowledge recycling:* This problem concerns the issue of not using previously gained knowledge in the maintenance work. A root cause analysis was recorded for all unplanned shutdowns that lasted more than one hour. An existing template was used and saved in a word file but the form was difficult to access, which resulted in old files that were rarely consulted or followed-up.

12) *Poor connections between teams:* This problem relates to the transfer of information between maintenance teams. When, an operator made a failure notice, it was categorized as either a mechanical or as an electrical fault. However, some failures are not only rooted in either a mechanical or an electrical problem and they can only be solved with electrical and mechanical technicians working together. The normal operation was that the work order was sent to one of the groups and, since, they cannot solve the problem, they direct it to the other group, and if they cannot find a solution they direct it back and so on. Both groups have different foremen and separate coffee rooms; thus, information sharing between the groups was not facilitated by the system or the organizational structure.

13) *Poor IT knowledge:* Computerized maintenance management systems (CMMS) are usually made to support the maintenance organization in their tasks. CMMS can be used as the IT-solution to collect data and support people in their work tasks and is used to manage the whole operative and strategic actions in the plant. However, this study showed that CMMSs were mainly used by maintenance organisations, even though; the operative staff completes a lot of maintenance actions. This lead to the situation that only part of the actual maintenance actions were recorded in the CMMS and therefor a great deal of data is not collected. Finally, when the operative and strategic decisions were made, they were based on the recorded data which was only showed a part of which occurred. Therefore, the strategic decisions were often not well grounded.

B. Data Quality in Maintenance

The perceived problems, presented in the previous section, were all rooted in different aspects of DQ. Table 2 gives an overview of the different problems identified during the studies, an also linkage between the problems and the phases in the maintenance process.

The data quality problems are listed in rows and they are correlated with maintenance phases in the columns. This table shows that the maintenance preparation phase has the greatest number of deficiencies. This shows that preparation and planning are necessary factors for improving DQ, but as in this case the lack of preparation affected the whole system negatively. A problem that showed up in all the maintenance phases was the lack of transparency in the system. If a focus was placed on usability, several of the other problems would most likely be reduced. Both planning and clearly understandable instructions, tasks and systems would increase the DQ level greatly.

TABLE 2.
IDENTIFIED DQ ISSUES AND THEIR LINKAGE TO MAINTENANCE PROCESS.

Data quality problem	Maintenance Management phase	Maintenance Support Planning phase	Maintenance Preparation phase	Maintenance Execution phase	Maintenance Assessment phase	Maintenance Improvement phase	Type of data quality problem
Multiple information sources in spare parts			√				Intrinsic
Lacking progress information for spare part orders		√	√				Contextual Accessibility
Manually inputted work orders varies in quality			√	√	√		Intrinsic Contextual
Information needed is not connected to work orders				√			Contextual Accessibility
Lack of information needed to follow up work					√	√	Intrinsic Contextual
Multiple failure notices			√	√	√		Intrinsic
Usability of systems	√	√	√	√	√	√	Representational
Information duplication and poor maintenance of information		√	√	√	√	√	Intrinsic
Performance indicators missing or difficult to obtain	√				√	√	Contextual
Bad information logistics between teams			√	√			Accessibility
Poor support for work planning by system		√	√				Representational Accessibility
Route cause analysis poorly utilized					√	√	Accessibility

VI. CONCLUSIONS

eMaintenance solutions should provide mechanisms that measure, manage and visualise the Quality of Service and DQ. The conducted case study indicates that DQ can be related to different features (e.g. usability, accuracy, and relevancy). These DQ features need to be visualised in order to be measured and managed by the eMaintenance solution and its user.

There exist some generic contributions, which deal with DQ, but the greatest need is to develop quality assurance mechanisms within eMaintenance solutions. These mechanisms should encompass all phases of the maintenance process; be able to manage DQ from different data sources and at various aggregation levels; to visualise DQ for the user; be able to adapt to user's context and the decision process; besides, criticality management. Furthermore, it is also important that there is an overarching plan for information management that covers a system's whole lifecycle

REFERENCES

[1] A. H. C. Tsang. "Strategic dimensions of maintenance management," *Journal of Quality in Maintenance Engineering*, vol. 8, no. 1, pp. 7-39, 2002.

[2] R. Karim. *A Service-Oriented Approach to eMaintenance of Complex Technical Systems*. Doctoral thesis, Luleå University of Technology, Luleå, Sweden, 2008.

[3] R. Karim. "eMaintenance industrial applications: Issues and challenges," *World Congress on Engineering Asset Management*, Cincinnati, Ohio, October 3, 2011, pp. 6.

[4] IEC 60300-3-14:*Dependability management - Part 3-14: Application guide - Maintenance and maintenance support.* International Elechtrotecnical Commission, Geneva, Switzerland, 2004.

[5] M. Kajko-Mattsson, R. Karim, and A. Mirijamdotter. Fundamentals of the eMaintenance Concept. Paper presented at the *1st Conference on eMaintenance*, Luleå, Sweden, 2010.

[6] E. Levrat, B. Iung & A. Crespo Marquez. "E-maintenance: review and conceptual framework," *Journal of Production Planning & Control*, vol. 19, no. 4, pp. 1–22, 2008.

[7] A. Muller, A. Crespo Marquez, and B. Iung. "On the concept of eMaintenance: Review and current research," *Reliability Engineering and System Safety,* vol. 93, no. 8, pp. 1165-1187, 2008.

[8] M. Koc and J. A. Lee. "A system framework for next generation EMaintenance Systems," In Proceedings from *2nd international Symposium on Environmentally Conscious Design and Inverse Manufacturing*, Tokyo, 2001

[9] A. Parida and U. Kumar. "Managing information is key to maintenance effectiveness," In Proceedings from *Intelligent Maintenance Systems (IMS) International Conference*, Arles, 2004.

[10] ISO 9000:2005, *Quality management systems, Fundamentals and vocabulary*, International Organization for Standardization, 2005.

[11] R. Y. Wang and D. M. Strong. "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems I,* Spring, vol. 12, no. 4, pp. 5-34, 1996.

[12] E. Adir, G. Shankaranarayanan and P. D. Berger. "Evaluating a model for cost-effective data quality management in a real-world CRM setting," *Decision Support Systems*, vol. 50, no. 1, 2010.

[13] ISO. ISO/TS 8000-100:2009, *Data quality — Part 100: Master data: Exchange of characteristic data: Overview*: International Organization for standardization, 2009.

[14] R. Price and G. A. Shanks. "Semiotic information quality framework: Development and comparative analysis," *Journal of Information Technology*, vol. 20, no. 2, pp. 88-102, 2005.

[15] Y. Wand and R. Y. Wang. "Anchoring data quality dimensions in ontological foundations," *Communications of the ACM*, vol. 39, no. 11, pp. 86-95, 1996.

[16] R. J. Price and G. Shanks. "A Semiotic Information Quality Framework", *Decision Support in an Uncertain and Complex World: The IFIP TC8/WG8.3 International Conference*, Tuscany, Italy, 2004.

[17] J. A. Harding, M. Shahbaz, Srinivas, and A. Kusiak. "Data mining in manufacturing: A review," *Journal of Manufacturing Science and Engineering*, vol. 128, no. 4, pp. 969-976, 2006.