Data Quality Assessment of Company's Maintenance Reporting: A Case Study

Manik Madhikermi[1], Sylvain Kubler[2], Andrea Buda[1] and Kary Främling[1]

[1]*Aalto University, School of Science P.O. Box 15400, FI-00076 Aalto, Finland*

[2]*University of Luxembourg, Interdisciplinary Centre for Security, Reliability & Trust, Luxembourg, Luxembourg*

Keywords: Data Quality, Multi-criteria Decision Making, Analytic Hierarchy Process, Decision Support Systems.

Abstract: Businesses are increasingly using their enterprise data for strategic decision-making activities. In fact, information, derived from data, has become one of the most important tools for businesses to gain competitive edge. Data quality assessment has become a hot topic in numerous sectors and considerable research has been carried out in this respect, although most of the existing frameworks often need to be adapted with respect to the use case needs and features. Within this context, this paper develops a methodology for assessing the quality of enterprises' daily maintenance reporting, relying both on an existing data quality framework and on a Multi-Criteria Decision Making (MCDM) technique. Our methodology is applied in cooperation with a Finnish multinational company in order to evaluate and rank different company sites/office branches (carrying out maintenance activities) according to the quality of their data reporting. Based on this evaluation, the industrial partner wants to establish new action plans for enhanced reporting practices.

1 INTRODUCTION

Data and Information quality[1] is one of the most competitive advantages for an organization in today's digital age (e.g., with the rapid evolution of Cloud Computing, Internet of Things – IoT, Big Data...) (Atzori et al., 2010). Companies are trying hard to find out relevant strategies to make their products (physical or virtual products) standout with respect to their competitors. In such environments, companies need to provide after sales services such as maintenance, and warranty services, in order to ensure that the delivered product is reliable and in full accordance with the customer requirements. Nonetheless, providing such services inevitably generate costs for businesses; within many industries, maintenance costs can account for up to 40% of the operational budget (Dunn, 1998). Some surveys indicate that one third of every dollar of maintenance costs is wasted due to inappropriate or unnecessary maintenance practices (Mobley, 2002). In fact, data quality practices (including maintenance reports) has a considerable impact on these costs since poor data quality impacts the downstream part of the maintenance process, and reciprocally, high data quality fosters enhanced business activities and decision making.

Data quality has been intensively studied over the last two decades, and various relevant frameworks for assessing data quality have since then emerged (Krogstie et al., 1995; Wang and Strong, 1996; Jarke and Vassiliou, 1997), and continue to emerge (Batini et al., 2009; Price and Shanks, 2009). Although most of the conceptual data quality frameworks can be applied regardless of the application area, they often require some tuning/adaptation to each use case needs and peculiarities, e.g. when dealing with healthcare, environmental, governmental, business, or still engineering applications (Berndt et al., 2001; Peabody et al., 2004). The present article is set within this context of 'existing framework adaptation', whose ultimate goal of our study is to assess company's maintenance reporting quality considering different office branches of a Finnish multinational Original Equipment Manufacturer (OEM). In light of the Multi-Criteria Decision Making (MCDM) nature of the problem (further described in Section 2), our study proposes to combine a conceptual data quality framework, namely Krogstie's framework (Krogstie et al.,

[1]The terms Data and Information are often used synonymously; in practice, managers differentiate information from data intuitively, and describe information as data that has been processed and enriched in some manner but, unless specified otherwise, this article will use "information" interchangeably with "data".

Madhikermi M., Kubler S., Buda A. and Främling K..
Data Quality Assessment of Company's Maintenance Reporting: A Case Study.
DOI: 10.5220/0005518401620172
In *Proceedings of 4th International Conference on Data Management Technologies and Applications* (DATA-2015), pages 162-172
ISBN: 978-989-758-103-8

1995), with a simple and effective MCDM technique aiming at aggregating the different data quality dimensions so as to come up with a ranking of the different company's sites in order of maintenance reporting quality.

To this end, section 2 introduces both the Krogstie's framework and to what extent it is adapted to our maintenance use case. Section 3 provides greater detail about the adaptation steps and its combination with the MCDM technique. Section 4 presents the use case results related to the OEM company, along with the conclusions.

2 DATA QUALITY FRAMEWORK AND ADAPTATION

Data quality is a well explored domain, in which many frameworks have emerged. One of the earlier framework was developed by Wang and Strong in (Wang and Strong, 1996), followed by many other scholars (Jarke and Vassiliou, 1997; Kahn et al., 2002; Batini et al., 2009; Price and Shanks, 2009). Despite differences in methods and contexts, yet they share a number of characteristics regarding their classifications of the quality dimensions (see e.g. the sixteen dimensions introduced by Wand and Strong). It is difficult to state in what respects one framework is better than another since data quality is commonly thought of as a multi-dimensional concept with varying attributed characteristics, which depend on the author's philosophical viewpoint, past experience, application domains, and so forth (Knight and Burn, 2005). Within this context, the scientific contribution of this paper is not to define a new data quality framework, but rather to apply and adapt a traditional one so as to cope with the company's needs, expectations and application features. Accordingly, section 2.1 provides a brief introduction of the considered framework, followed by section 2.2 that details to which extent this framework is used/extended to our needs.

2.1 Reference Data Quality Framework

The data quality framework considered in our study is the one defined by Krogstie et al. (Krogstie et al., 1995), which is an extension of the framework defined by (Lindland et al., 1994). The different concepts and relationships of the Krogstie's framework are illustrated in Figure 1, which consists of:

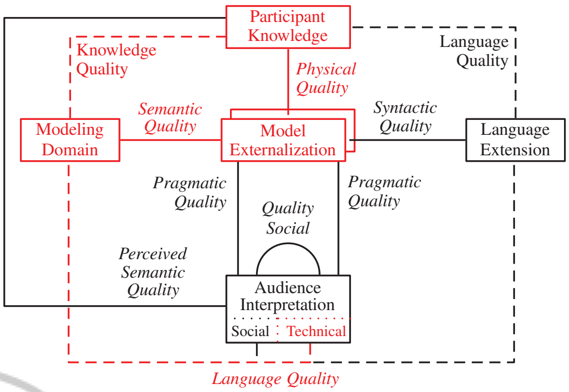

Figure 1: Krogstie's data quality framework.

- *Physical Quality:* about externalizability (i.e., the knowledge of some social actors have been externalized by the use of a conceptual modeling language) and internalizability (i.e., the externalized model is persistent and available enabling participants to make sense of it);
- *Syntactic Quality:* correspondence between the model and the language extension of the language in which the model is written;
- *Semantic Quality:* correspondence between the model and the domain, where the domain is considered as the ideal knowledge about the situation to be modeled. Krogstie's framework contains two semantic goals: *Validity* and *Completeness*;
- *Perceived Semantic Quality:* correspondence between the actor interpretation of a model and his/her current knowledge of the domain. In line with the semantic quality, two goals are defined by the authors: *Perceived Validity* and *Perceived Completeness*;
- *Pragmatic Quality:* correspondence between the model and the "Audience Interpretation" of it (*cf.* Figure 1);
- *Social Quality:* about people "agreement";
- *Knowledge Quality:* from a pure standpoint of social construction, and as stated by Krogstie et al., it is difficult to talk about the quality of explicit knowledge. On the other hand, within certain areas such as mathematics, what is regarded as 'true' is comparatively stable, and it is intersubjectively agreed that certain people have more valid knowledge of an area than others. The 'quality' of the participant knowledge can thus be expressed by the relationships between the audience knowledge and the domain.
- *Language Quality:* appears as means for model quality in the framework. Krogstie et al. have regrouped factors from earlier discussions on language quality as follows:

Table 1: Criteria and its sub-criteria description related to the data quality dimensions.

Criteria	Sub-Criteria	Description	Type
Believability (C_B)	Length of Work Description (C_{B1})	Length of the work description related to a work order.	$I^c_{\text{avg}}(i)$
	Work Log Variation (C_{B2})	Work Description variation among the different operator reports	$I^c_{\text{sim}}(i)$
	Technician Log Variation (C_{B3})	Technical log variation among the different operator reports	$I^c_{\text{sim}}(i)$
Completeness (C_C)	Asset Location reported (C_{C1})	Location of asset within product where maintenance has been done.	$I^c_{\text{sim}}(i)$
	Description reported (C_{C2})	Description of work to be done in particular maintenance work.	$I^c_{\text{sim}}(i)$
	Actual Finish Date reported (C_{C3})	Actual Finish date and time of work completed.	$I^c_{\text{sim}}(i)$
	Target Start Date reported (C_{C4})	Targeted start date of the maintenance work.	$I^c_{\text{sim}}(i)$
	Target Finish Date reported (C_{C5})	Targeted finish date of the maintenance work.	$I^c_{\text{sim}}(i)$
	DLC Code reported (C_{C6})	Actual location of the defect within product.	$I^c_{\text{sim}}(i)$
	Schedule Start Date reported (C_{C7})	Scheduled start date of the maintenance work.	$I^c_{\text{sim}}(i)$
	Schedule Finish Date reported (C_{C8})	Scheduled Finish date of the maintenance work.	$I^c_{\text{sim}}(i)$
Timeliness (C_T)		This is average delay of reporting on individual site	$I^c_{\text{avg}}(i)$

– Domain Appropriateness;
– Participant Knowledge Appropriateness;
– Technical Actor Interpretation Enhancement.

2.2 Krogstie's Framework Adaptation

Given the above definitions, and based on the OEM company's requirements, three key concepts/relationships and one assumption lay the groundwork of our study for Krogstie's framework adaptation. First, the study assumption is that the *Physical Quality* (*cf.* Figure 1), and particularly the externalized model, is 100% persistent and available, thus enabling participants to make sense of it. Indeed, the OEM company designed its own maintenance models, report templates, databases, *etc.*, and is not willing (at a first stage) to assess/study how persistent their implementations are compared with the initial expert statements, expressed knowledge, *etc.* The OEM company then expressed requirements regarding three of the Krogstie's framework concepts/relationships, namely:

1. *Semantic Quality:* one of the OEM company's requirement matches – to a certain extent – with the semantic quality dimension since the company would like to know to which extent the service data reported by each operator (on each site) can be trusted, or more exactly can be considered as "true", "real" and "credible", in order to carry out the planning activities. This is referred to as the "Believability" criterion (C_B) in this paper, whose various facets of the Believability are formalized in the form of sub-criteria (or Believability quality indicators) denoted by $\{C_{B1}..C_{B3}\}$ in Table 1;
2. *Language Quality:* one of the OEM company's requirement matches – to a certain extent – with the language quality dimension since the company would like to know to which extent the service data reported by each operator is complete, or is of sufficient depth and breadth for the task at hand (Wang and Strong, 1996). To put it another way, this criterion, referred to as Completeness (C_C), reflects the level of details reported by each operator with regard to each report field that needs to be entered (in accordance with the company's business logic) in the report. Similarly to C_B, the facets of Completeness are denoted $\{C_{C1} \dots C_{C8}\}$ (see Table 1);
3. *Knowledge Quality:* one of the OEM company's requirement matches – to a certain extent – with the semantic quality dimension since the company would like to know to which extent the service data reported by each operator is sufficiently "up to date", which is depending on the time difference between the maintenance work and the work reporting. This criterion, referred to as Timeliness C_T, is based on the assumption that the longer the time spent to submit the report, the lesser the quality of the reporting (operator are likely to forget key details of the maintenance task over time). No sub-criterion is defined for this dimension, as shown in Table 1 (C_T);

In order to ease the understanding of these three data quality dimensions, and associated sub-criteria, we propose to illustrate through Figure 2 the different stages that compose our adapted framework. This figure highlights that maintenance operators carry out maintenance work/tasks on each OEM site (sites denoted by Site 1... Site z) and generate multiple reports. A zoom on reports from Site 1 and n is proposed in Figure 2 so as to compare both sets of reports based on the criteria defined in Table 1. It allows for an understanding of when a report, or field content, impacts positively on the company's maintenance reporting quality, and when it does impact negatively (see "smileys" and associated explanation in Figure 2).

In this paper, a simple and effective MCDM technique is used as support of the arithmetic framework to handle the integration/aggregation of the various

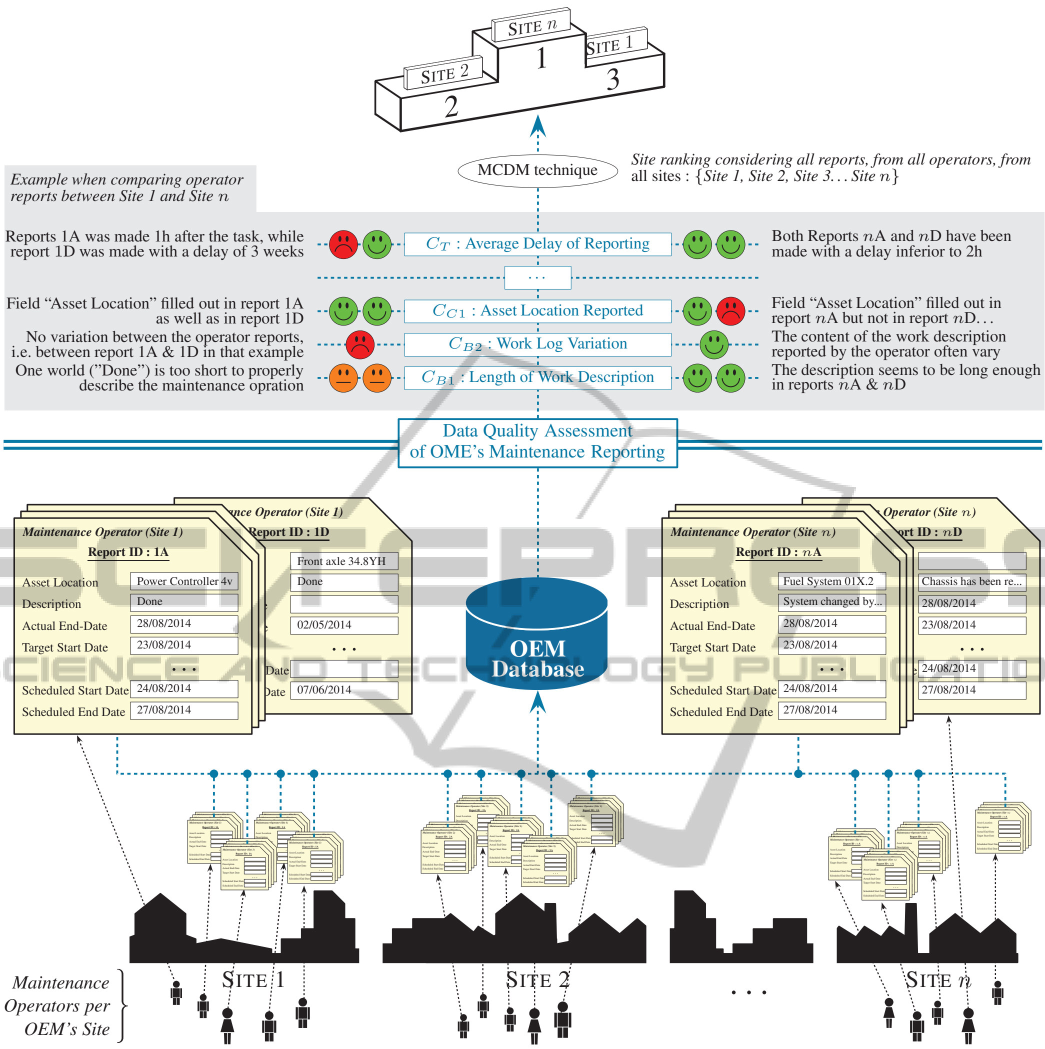

Figure 2: Stages composing the maintenance reporting quality assessment framework.

criteria preferences, report contents, *etc.* as emphasized in Figure 2 (see the podium that is the result of the "MCDM technique"). The reason of using a MCDM technique is threefold:

- the human brain is not reliable for decision-making when there are many factors/criteria to consider simultaneously, which is even more true when the problem is structured in several layers (i.e., objective depending on several criteria, which themselves can be declined into sub-criteria. . .), as it is the case in our use case;
- MCDM techniques help reasoning about interdependencies among criteria, alternatives, *etc.*, which inevitably results in better decision-making, or assessment outcomes;
- Experts from the OEM company can easily re-use and adapt the MCDM parameters as they see fit (e.g., criteria preferences, integration of new data quality dimensions);

There is a number of MCDM techniques in the literature such as AHP (analytic hierarchy process), ANP (analytic network process), TOPSIS (technique for order preference by similarity to ideal situation), ELECTRE to solve MCDM problems (Figueira et al., 2005). In our study, we do use AHP (Saaty, 1996) for the reason that it is very simple and effective technique to integrate expert opinions and requirements. For instance, decision makers use linguistic variables in AHP rather than expressing their judgments in the form of exact numeric values; adding that AHP does not involve complex mathematics. These characteristics are probably the main reasons for the success of this technique, which is the second most used MCDM

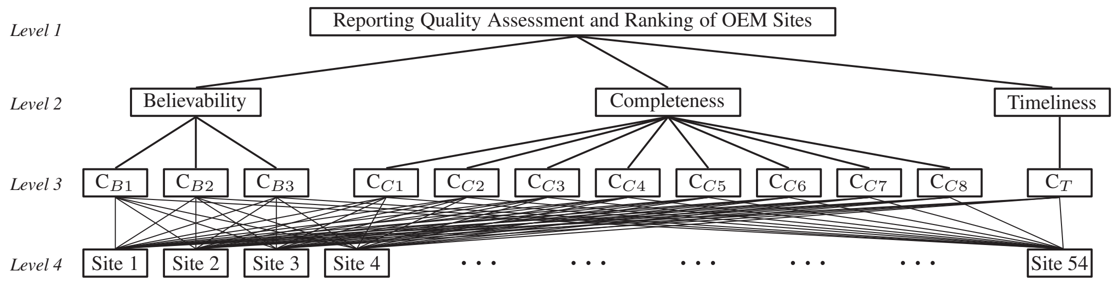

Figure 3: AHP structure related to the maintenance reporting quality assessment problem.

methods according to a recent survey[2] (Mardani et al., 2015). Nonetheless, it is important to note that there are no better or worse techniques, but some techniques are better suited to particular decision problems than others (Zheng et al., 2012); for instance, AHP only deals with linear preferences (this is the case in our study), not with contextual preferences where the value of one or several criteria may affect the importance or utility of other criteria (Främling, 1996).

3 DATA REPORTING ASSESSMENT

AHP, originally introduced by (Saaty, 1996), has the advantage of organizing critical aspects of the problem in a manner similar to that used by the human brain in structuring the knowledge, i.e. in a hierarchical structure of different levels consisting of the overall goal, the criteria and sub-criteria, as well as the alternatives. In this regard, our MCDM ranking problem is broken down into the hierarchical structure depicted in Figure 3, and particularly in four distinct levels:

- *Level 1:* the overall goal of the study is to rank the different OEM company sites in terms of maintenance reporting quality;
- *Levels 2 and 3:* the set of data quality dimensions, and sub-criteria, used to assess the maintenance reporting quality (derived from Krogstie's framework and listed in Table 1);
- *Level 4* the alternatives that are the OEM company sites;

Given this hierarchy, AHP does perform the following computation steps for identifying the final ranking of the alternatives with respect to the overall goal:

1. Compare each element in the corresponding level and calibrate them on the numerical scale. This requires $\frac{(n-1)}{2}$ pairwise comparisons, where n is the number of elements with the consideration that diagonal elements are equal to "1" and the other elements will be simply the reciprocal of the earlier comparisons;
2. Perform calculation to find the maximum eigen value, consistency index (CI), consistency ratio (CR), and normalized values for each criteria/alternatives;
3. If the computed eigen value, CI and CR are satisfactory, then decision/ranking is done based on the normalized values.

Stages 1 and 2 are detailed in sections 3.1 and 3.2, which respectively deal with expert preference-based pairwise comparisons and ratio scale-based pairwise comparisons (Saaty, 1990), and Stage 3 is described in section 3.3. In order to make the understanding easier, a scenario is considered throughout section 3, whose parts are preceded by the symbol "➪".

3.1 Pairwise Comparison based on Expert Preferences

This section details how a decision maker evaluates the importance of one criterion (or sub-criterion) with respect to the others. To this end, OEM experts perform pairwise comparisons among criteria, as formalized with P_{C} in Eq. 1, with m the number of criteria at a specific hierarchy level and from a same "parent criterion", e.g. $m = 3$ at level 2 of the AHP structure (i.e., $m = |\{\mathrm{C}_B, \mathrm{C}_C, \mathrm{C}_T\}|$), $m = 3$ at level 3 with regard to the parent criterion 'Believability' (i.e., $m = |\{\mathrm{C}_{B1}, \mathrm{C}_{B2}, \mathrm{C}_{B3}\}|$), $m = 8$ at level 3 with regard to the parent criterion 'Completeness', *etc.* The expert evaluation is carried out based on the 1- to 9-point Saaty's scale: $\{1,3,5,7,9\}$; $w_{ij} = 1$ meaning that C_i and C_j are of equal importance and $w_{ij} = 9$ meaning

[2] Frequency of application being 15.82% for AHP, while Hybrid Fuzzy MCDM (1st position) are applied with a frequency of 19.89% and Fuzzy AHP (3rd position) with a frequency of 9.53%.

$$\begin{array}{cc} & \begin{array}{cccccccc} \mathrm{C}_{C1} & \mathrm{C}_{C2} & \mathrm{C}_{C3} & \mathrm{C}_{C4} & \mathrm{C}_{C5} & \mathrm{C}_{C6} & \mathrm{C}_{C7} & \mathrm{C}_{C8} \end{array} \\ \begin{array}{c} \mathrm{C}_{C1} \\ \mathrm{C}_{C2} \\ \mathrm{C}_{C3} \\ \mathrm{C}_{C4} \\ \mathrm{C}_{C5} \\ \mathrm{C}_{C6} \\ \mathrm{C}_{C7} \\ \mathrm{C}_{C8} \end{array} & \begin{bmatrix} 1 & 3 & 1 & 3 & 7 & 3 & 9 & 3 \\ 1/3 & 1 & 1/3 & 3 & 5 & 3 & 5 & 3 \\ 1 & 3 & 1 & 3 & 5 & 3 & 5 & 3 \\ 1/3 & 1/3 & 1/3 & 1 & 5 & 3 & 5 & 1 \\ 1/7 & 1/5 & 1/5 & 1/5 & 1 & 1/3 & 3 & 5 \\ 1/3 & 1/3 & 1/3 & 1/3 & 3 & 1 & 5 & 1/3 \\ 1/9 & 1/5 & 1/5 & 1/5 & 1/3 & 1/5 & 1 & 1/5 \\ 1/3 & 1/3 & 1/3 & 1 & 1/5 & 3 & 5 & 1 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} W_{C_{C1}} \\ W_{C_{C2}} \\ W_{C_{C3}} \\ W_{C_{C4}} \\ W_{C_{C5}} \\ W_{C_{C6}} \\ W_{C_{C7}} \\ W_{C_{C8}} \end{array} \begin{bmatrix} 0.240 \\ 0.165 \\ 0.191 \\ 0.128 \\ 0.081 \\ 0.085 \\ 0.019 \\ 0.089 \end{bmatrix} \quad (4)$$

that C_i is strongly favored over C_j.

$$P_C = \begin{array}{cc} & \begin{array}{ccc} \mathrm{C}_1 & \dots & \mathrm{C}_m \end{array} \\ \begin{array}{c} \mathrm{C}_1 \\ \vdots \\ \mathrm{C}_m \end{array} & \begin{bmatrix} w_{11} & \dots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \dots & w_{mm} \end{bmatrix} \end{array} \quad (1)$$

The computation of the normalized eigenvector of P_{C} then enables to turn qualitative data into crisp ratios. Although several approaches exist in the literature for normalized eigenvector compution, the Simple Additive Weighting (SAW) method (Tzeng and Huang, 2011) is used in our study, as formalized in Eq. 2.

$$W_i = \frac{\sum_{j=1}^{m} w_{ij}}{\sum_{k=1}^{m}\sum_{j=1}^{m} w_{kj}}, \quad w_{ji} = \begin{cases} 1 & i = j \\ \frac{1}{w_{ij}} & i \neq j \end{cases} \quad (2)$$
$$\mathcal{W} = [W_{C_1}, \dots, W_{C_i}, \dots, W_{C_m}]$$

Finally, a P_{C} matrix is characterized as consistent if, and only if:

$$\mathrm{w}_{ij} = \mathrm{w}_{ik} \times \mathrm{w}_{kj} \ \forall i,k \in \mathcal{N} | i \neq k; j \in \mathcal{N} - \{i,k\}$$

However it is often hard to fulfill such a pre-requisite when dealing with real expert preferences, which is all the more true when the number of criteria to be compared increases. Consistency of any matrix is calculated through the Consistency Ratio (CR), as given in Eq. 3, where RI is the Consistency index of a pairwise matrix generated Randomly (Saaty, 1980).

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \quad (3)$$

➪ In our case, pairwise comparisons are filled out with the OEM's executive officer. Eq. 5 provides insight into the expert specifications regarding criteria at Level 2 of the AHP structure. The computed normalized eigenvector highlights that the officer judges all criteria at this level of equal importance.

$$\begin{array}{cc} & \begin{array}{ccc} \mathrm{C}_B & \mathrm{C}_C & \mathrm{C}_T \end{array} \\ \begin{array}{c} \mathrm{C}_B \\ \mathrm{C}_C \\ \mathrm{C}_T \end{array} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} W_{C_B} \\ W_{C_C} \\ W_{C_T} \end{array} \begin{bmatrix} 0.33 \\ 0.33 \\ 0.33 \end{bmatrix} \quad (5)$$

$\boxed{\mathrm{CI}=0; \mathrm{CR}=0}$

Eq. 6 shows the pairwise comparisons carried out at Level 3 of the AHP structure, with regard to the parent criterion 'Believability' (to facilitate understanding, the calculation of the normalized eigenvector value $\mathrm{W}_{C_{B1}}$ is detailed in Eq. 7). The eigenvector values (*cf.* Eq. 6) highlight that the officer judges the "Length of Work Description" slightly more important (or critical) in the maintenance reporting quality than the "Work Log Variation" (C_{B1}), and far more important than the "Technician Log Variation" (C_{B3}).

$$\begin{array}{cc} & \begin{array}{ccc} \mathrm{C}_{B1} & \mathrm{C}_{B2} & \mathrm{C}_{B3} \end{array} \\ \begin{array}{c} \mathrm{C}_{B1} \\ \mathrm{C}_{B2} \\ \mathrm{C}_{B3} \end{array} & \begin{bmatrix} 1 & 3 & 5 \\ \frac{1}{3} & 1 & 5 \\ \frac{1}{5} & \frac{1}{5} & 1 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} W_{C_{B1}} \\ W_{C_{B2}} \\ W_{C_{B3}} \end{array} \begin{bmatrix} 0.54 \\ 0.38 \\ 0.08 \end{bmatrix} \quad (6)$$

$\boxed{\mathrm{CI}=0.168; \mathrm{CR}=0.289}$

$$\mathrm{W}_{C_{B1}} = \frac{1+3+5}{1+3+5+\frac{1}{3}+1+5+\frac{1}{5}+\frac{1}{5}+1} \quad (7)$$
$$= \frac{9}{16.74} = 0.54$$

Similarly, the experts carry out pairwise comparisons in Eq. 4 considering the sub-criteria of 'Completeness' (i.e., C_{C1} to C_{C8}); $W_{C_{C1}}$ is the most important sub-criteria, followed by $W_{C_{C3}}$ and $W_{C_{C2}}$ respectively. Regarding C_T, there is no pairwise comparison be performed since no sub-criterion has been defined.

The pairwise comparison approach introduced in this section allows for taking into consideration expert know-how and judgments, and to turn them into crisp ratios. However, pairwise comparison evaluation is not always based on expert elicitation, sometimes them is necessary to take into consideration monitoring system parameters such as: how many times the field "DLC Code reported" (C_{C6}) has been left empty in the maintenance reports on Site *i* compared with the other Sites. In this case, Saaty introduced the concept of 'relative scale' or 'pairwise comparison as ratios" (Saaty, 1990), which allows for considering various types of data and metrics. Section 3.2 provides greater detail about the types of data and metrics that underly our pairwise comparisons as ratios that mostly concern pairwise comparisons among alternatives (i.e., level 4 of the AHP structure) with respect to a each criterion taking place at the upper level (i.e., at Level 3).

3.2 Pairwise Comparison as Ratios

Pairwise Comparison as ratios is a tool that allows for comparing criteria (or alternatives with respect to criteria) based upon a relative scale rather than using preference scales (e.g., the 1- to 9-point Saaty's scale). Eq. 8 provides insight into the pairwise comparison as ratio matrix considering the set of alternatives A_i (i.e., i referring to a OEM site), with $\mathrm{I}^c_x(i)$ the digital indicator (or metric) that enables us to quantitatively assess the alternative A_i with respect to the monitored system parameter c (i.e., with respect to criteria defined at Level 3), and x referring to the fact that several digital indicators can be used according to the monitored system parameter/criterion c, as will be discussed below. Note that the normalized eigenvector values of the pairwise comparison as ratios with respect to criterion c are denoted by $W_{A_i^c}$ in Eq. 8.

$$\begin{array}{c} \\ \mathrm{A}_1 \\ \mathrm{A}_2 \\ \vdots \\ \mathrm{A}_z \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{A}_1 & \mathrm{A}_2 & \dots & \mathrm{A}_z \end{array} \\ \begin{bmatrix} 1 & \frac{\mathrm{I}^c_x(1)}{\mathrm{I}^c_x(2)} & \dots & \frac{\mathrm{I}^c_x(1)}{\mathrm{I}^c_x(z)} \\ \frac{\mathrm{I}^c_x(2)}{\mathrm{I}^c_x(1)} & 1 & \dots & \frac{\mathrm{I}^c_x(1)}{\mathrm{I}^c_x(z)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{I}^c_x(z)}{\mathrm{I}^c_x(1)} & \frac{\mathrm{I}^c_x(z)}{\mathrm{I}^c_x(2)} & \dots & 1 \end{bmatrix} \end{array} \Rightarrow \begin{bmatrix} W_{A_1^c} \\ W_{A_2^c} \\ \vdots \\ W_{A_z^c} \end{bmatrix} \tag{8}$$

Two digital indicators $\mathrm{I}^c_x(i)$ are defined:

- $\mathrm{I}^c_{\mathrm{sim}}(i)$ (Empty Indicator – Eq. 9): used to calculate the number of times a "field" was left empty in reports carried out on Site i, with k the total number of reports performed on Site i:

$$\mathrm{I}^c_{\mathrm{sim}}(i) = \frac{\text{Number of empty fields on Site } i}{k} \tag{9}$$

➪ Let us consider the example of pairwise comparison as ratios with regard to C_{C6} and Site 1 and 2. On Site 1, 76 maintenance reports have been carried out and 45 of these reports contain the DLC code (meaning that 59% of all the reports contain the requested information, see Eq. 10), while on Site 2 only 44% of the reports contain the requested information (see Eq. 11).

$$\mathrm{I}^{\mathrm{C}_{C6}}_{\mathrm{sim}}(1) = \frac{45}{76} = 59\% \tag{10}$$

$$\mathrm{I}^{\mathrm{C}_{C6}}_{\mathrm{sim}}(2) = \frac{49}{88} = 44\% \tag{11}$$

The pairwise comparison as ratios is then computed using all $\mathrm{I}^c_x(i)$ indicators and considering all alternatives (i.e., the 54 sites). Eq. 12 provides insight into such pairwise comparison as ratios with respect to C_{C6}, in which $\mathrm{I}^{\mathrm{C}_{C6}}_{\mathrm{sim}}(1)$ and $\mathrm{I}^{\mathrm{C}_{C6}}_{\mathrm{sim}}(2)$ (computed above) are used.

$$\begin{array}{c} \\ \mathrm{A}_1 \\ \mathrm{A}_2 \\ \vdots \\ \mathrm{A}_{54} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{A}_1 & \mathrm{A}_2 & \dots & \mathrm{A}_{54} \end{array} \\ \begin{bmatrix} 1 & \frac{59}{44} & \dots & 0.15 \\ \frac{44}{59} & 1 & \dots & 0.67 \\ \vdots & \vdots & \ddots & \vdots \\ 6.64 & 1.50 & \dots & 1 \end{bmatrix} \end{array} \Rightarrow \begin{array}{c} W_{A_1}^{\mathrm{C}_{C6}} \\ W_{A_2}^{\mathrm{C}_{C6}} \\ \vdots \\ W_{A_{54}}^{\mathrm{C}_{C6}} \end{array} \begin{bmatrix} 0.187 \\ 0.002 \\ \vdots \\ 3\mathrm{E}\text{-}06 \end{bmatrix} \tag{12}$$

- $\mathrm{I}^c_{\mathrm{avg}}(i)$ (Average Indicator – Eq. 13): used to calculate the average delays for reporting the maintenance reports per site (i.e., regarding C_T) or the average length of work description (i.e., C_{B1}) per site. Mathematically, $\mathrm{I}^c_{\mathrm{avg}}(i)$ is computed based on Eq. 13, where q is either the reporting delay value or the description length value of one of the k reports carried out on Site i.

$$\mathrm{I}^c_{\mathrm{avg}}(i) = \frac{\sum_{q=1}^{k} q}{k} \tag{13}$$

➪ Let us assume that 4 maintenance reports have been carried out on Site 1, and that the work description length is equal to 44, 5, 13 and 101 respectively. In that case, the average indicator with regard to C_{B1} and Site 1 will be equal to 40.75 (see Eq. 14). Similarly to Eq. 12, the pairwise comparison as ratios is computed considering all $\mathrm{I}^c_x(i)$ indicators and all alternatives. The final matrix is not presented here due to the similarity with the one presented in Eq. 12.

$$\mathrm{I}^{\mathrm{C}_{B1}}_{\mathrm{avg}}(1) = \frac{44+5+13+101}{4} = 40.75 \tag{14}$$

Note that we highlighted in Table 1 (see last column) what indicators – $\mathrm{I}^c_{\mathrm{sim}}(i)$ or $\mathrm{I}^c_{\mathrm{avg}}(i)$) – is used with regard to each criterion.

3.3 Alternative Ranking

Figure 4 sums up all variables and related weights computed in the previous sections. It is now necessary to aggregate the different weights in order to converge towards a final ranking of the alternatives/sites. To this end, the global weight of each alternative with respect to all criteria C_x is computed based on Eq. 15.

$$GW_{A_i}^{\mathrm{C}_x} = W_{A_i}^{\mathrm{C}_x} \times W_{\mathrm{C}_x} \times W_{\mathrm{C}_{x(parent)}} \tag{15}$$

Let us apply this formula in Eq. 16 considering alternative A1 (i.e., Site 1) and criterion C_{C6}, whose

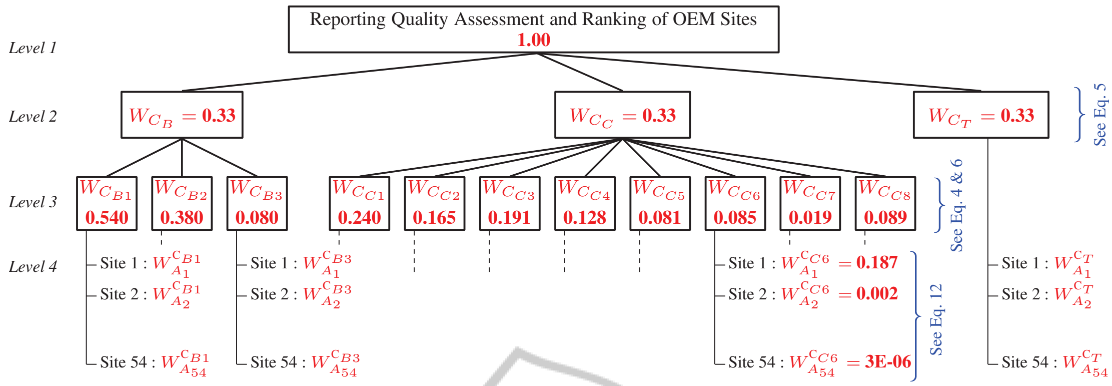

Figure 4: AHP structure and associated weights.

Table 2: Global Weight Computation of all Alternatives with respect to all Criteria.

	C_{B1}	C_{B2}	C_{B3}	$\sum C_{Bx}$	C_{C1}	C_{C2}	C_{C3}	C_{C4}	C_{C5}	C_{C6}	C_{C7}	C_{C8}	$\sum C_{Cx}$	$\sum C_T$
Site 1	$GW_{A_1}^{C_{B1}}$	…	$GW_{A_1}^{C_{B3}}$	$\sum_{x=\{1..3\}}\left(GW_{A_1}^{C_{Bx}}\right)$	…	…	…	…	…	$GW_{A_1}^{C_{C6}}$	…	…	$\sum_{x=\{1..8\}}\left(GW_{A_1}^{C_{Cx}}\right)$	$GW_{A_1}^{C_T}$
Site 2	$GW_{A_2}^{C_{B1}}$	…	$GW_{A_2}^{C_{B3}}$	$\sum_{x=\{1..3\}}\left(GW_{A_2}^{C_{Bx}}\right)$	…	…	…	…	…	$GW_{A_2}^{C_{C6}}$	…	…	$\sum_{x=\{1..8\}}\left(GW_{A_2}^{C_{Cx}}\right)$	$GW_{A_2}^{C_T}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Site 54	$GW_{A_{54}}^{C_{B1}}$	…	$GW_{A_{54}}^{C_{B3}}$	$\sum_{x=\{1..3\}}\left(GW_{A_{54}}^{C_{Bx}}\right)$	…	…	…	…	…	$GW_{A_{54}}^{C_{C6}}$	…	…	$\sum_{x=\{1..8\}}\left(GW_{A_{54}}^{C_{Cx}}\right)$	$GW_{A_{54}}^{C_T}$

"parent criterion" is logically C_C.

$$
\begin{aligned}
GW_{A_1}^{C_{C6}} &= W_{A_1}^{C_{C6}} \times W_{C_{C6}} \times W_{C_C} \\
&= 0.187 \times 0.085 \times 0.333 \\
&= 0.053
\end{aligned}
\tag{16}
$$

The global weight related to each alternative is then computed as summarized in Table 2. It is thus possible to aggregate those global weights per "parent criterion", i.e. regarding Believability (C_B) Completeness (C_C) and Timeliness (C_T) as formalized in the columns detoned by $\sum C_{Bx}$, $\sum C_{Cx}$ and $\sum C_T$ in Table 2.

We do not further detail the calculations, we rather provide (in Table 3) the final alternative/site ranking with regard to each "parent criterion"; e.g., Site 1 is ranked 17^{th} out of the 54^{th} sites in terms of 'Believability', 3^{rd} out of the 54^{th} sites in terms of 'Completeness', and 2^{nd} in terms of 'Timeliness'. Based on these first results, first conclusions can be drawn: Figure 5 provides a comparison view (using a spider chart) among different alternatives/sites (we voluntary did not include the 54 alternatives for clarity purposes) that helps us to see how good each company's site is with regard to each data quality dimension. Note that in this case, the wider the shape (e.g., Site 11 and 32 have the widest/biggest shapes), the better the company's site.

In order to obtain the final ranking of the alternatives, i.e. aggregating all alternative global weights into a single and final score, it is necessary to sum $\sum C_{Bx}$, $\sum C_{Cx}$ and $\sum C_T$ regarding each alternative/site. Such results are presented and discussed in section 4.

Table 3: Site ranking with respect to each data quality dimension (i.e., parent criteria).

	Believability	Completeness	Timeliness
Site 1	30^{th}	3^{rd}	2^{nd}
Site 2	4^{th}	15^{th}	27^{th}
Site 3	7^{th}	37^{th}	31^{st}
⋮	⋮	⋮	⋮
Site 11	**33^{rd}**	**7^{th}**	**1^{st}**
⋮	⋮	⋮	⋮
Site 32	**2^{nd}**	**4^{th}**	**18^{th}**
⋮	⋮	⋮	⋮
Site 37	**46^{th}**	**52^{th}**	**37^{th}**
⋮	⋮	⋮	⋮
Site 47	**19^{th}**	**35^{th}**	**31^{th}**
⋮	⋮	⋮	⋮

4 USE CASE RESULTS

This section presents the results of one experiment of the maintenance reporting quality assessment.

In practice, our tool has been developed with Matlab, which enables the executive officer to assess, at a given point in time, the quality of the different company's sites considering historical data/reports. The assessment period can be adjusted by the officer as he/she sees fit (e.g., to assess/compare sites over the previous days, weeks or months). The user interface (UI) provides the executive officer with the possibil-

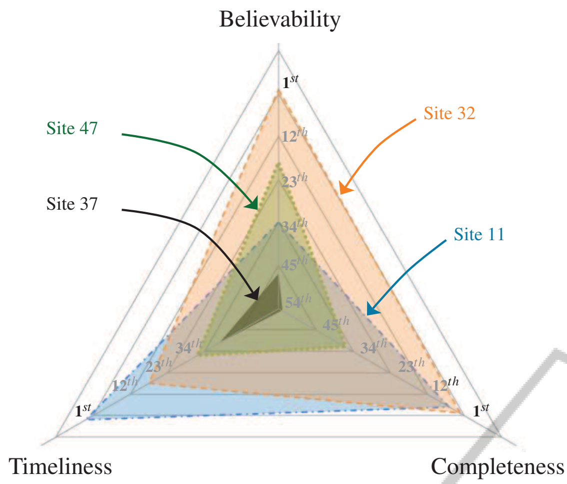

Figure 5: Comparison of sites 11, 32, 37 and 47.

ity to modify his/her preferences regarding the "pairwise comparison based on expert preferences". For example, if for some reasons he/she wants to give further importance to the "Completeness" dimension over Believability and Timeliness. Considering the pairwise comparison as ratios, such rations are computed by performing SQL queries against the OEM's information system that contains the maintenance reports (*cf.* Figure 2).

Based upon the executive officer preferences (the ones specified throughout section 3), the histogram in Figure 6 gives insight into the maintenance reporting quality assessment results: *x*-axis referring to the 54 sites, *y*-axis giving the quality maintenance reporting quality score. In total (considering all reports, from all sites), 275.585 reports have been processed and analyzed. The histogram shows that some quality scores dropped below "0"; the reason being that a penalty score has been introduced when a report field was left empty[3]. The histogram thus provides the overall ranking: Site 11 has the better quality score, followed by Site 1, Site 18...; Site 15 has the lowest quality score. Although the histogram does not provide enough information to identify the reasons for a good or non-standard reporting, it nonetheless provides first insights into qualitative results that may help to understand some of the reasons (e.g., a lack of training, insufficient manpower, ...). These results also offer the opportunity to identify and understand the good reporting practices from the best sites so as to learn and apply those practices on the less performant sites. Another action from the executive officer perspective is to cluster the sites based on reporting quality, thus enabling easier implementation of corrective actions driven by the clustering.

Again, let us remember that the executive officer has the possibility to customize his/her own UI dashboard by selecting different views, e.g. the histogram view (Figure 6), the spider chart view (Figure 5), *etc.*, each of them providing more or less detailed and aggregated information (the level of aggregation of the results varies depending upon the selected view).

5 CONCLUSIONS

In recent years, implementation of effective maintenance strategies proved to be a significant source for financial savings and enhanced productivity. At the heart of those strategies is the quality of data that includes, among other things, maintenance reporting activities. Indeed, maintenance data has directs impact on other company activities such as on:

- *after-sales services:* the quality of maintenance reports makes it possible to assess the maintenance work, thus helping to reach a higher quality after-sales services;
- *on the design of future generations of products:* processing and analyzing 'relevant' maintenance reports help to better understand how the products from the company behave throughout their product lifecycle, thus helping to enhance the design of the next product generations (Främling et al., 2013);
- *predictive maintenance strategies:* providing real-time and remote predictive maintenance is becoming a very promising area in the so-called IoT (Buda et al., 2015), whose objective is to provide systems with the capability to discover and process real-time data and contexts so as to make pro-active decisions (e.g., to self-adapt the system before a possible failure). Although real-time data is of the utmost importance in the predictive maintenance process, combining such data with historical maintenance reporting data (regarding a specific product item) has the potential to generate new knowledge, thus leading to more effective and product-centric decisions;
- *government regulation compliance:* in some domains, it is mandatory to comply with government regulations (e.g., in automotive, avionics, or healthcare domains). In this respect, assessing the quality of maintenance reporting can prevent the company from having regulation non-compliance

[3] Although other penalty strategies could be applied, we propose as a first step to define the penalty as $(-1 \times K)$ with K the criterion importance (signifying that the higher the criterion importance, the higher the penalty score for not having filled out the report field)

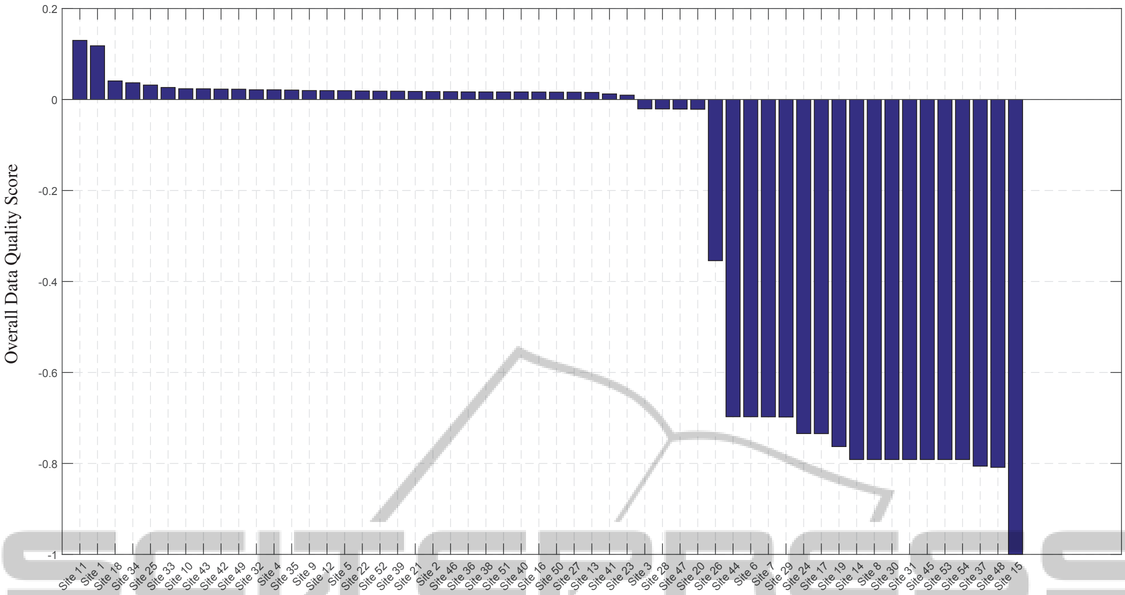

Figure 6: Site ranking according to the maintenance reporting quality assessment study.

issues, e.g. by carefully following the data quality on each company's site and identifying when the quality is too poor, or when a key data quality dimension is not of sufficient quality;

Given the above statements, a methodology for assessing the quality of enterprises' daily maintenance reporting is developed in this paper, which relies, on the one hand, on the Krogstie's data quality framework and, on the other hand, on a simple arithmetic MCDM framework (AHP) in order to handle the aggregation of the expert preferences, application features, *etc.* (the reason for combining both techniques being given in sections 2 and 3). An important aspect of our methodology, and adapted framework, is that this framework can further be extended in two respects:

- *Data quality framework extension:* as highlighted in Figure 1, only a few concepts and relationships from the Krogstie's framework were considered (semantic quality, knowledge quality. . .), which is mainly due to the company's expectations and needs. Accordingly, the framework can be further extended considering the other concepts/relationships (not used yet) such as Language Quality (e.g., for domain appropriateness, participant knowledge appropriateness. . .), Syntactic Quality (e.g., for syntactical correctness purposes, meaning that all statements in the model are according to the syntax of the language), and so forth;
- *AHP structure extension:* as described in section 2.2, a first set of criteria and sub-criteria have been considered, but further data quality dimensions can easily be added to the overall AHP structure (see Figure 3).

Our maintenance reporting quality assessment framework has been developed and applied in cooperation with a Finnish OEM company in order to evaluate and rank 54 office branches, which are spread in different countries. Based on this initial evaluation (*cf.* section 4), the OEM partner has since then established adapted action plans for enhanced reporting practices, and is now interested in extending this initial framework.

ACKNOWLEDGEMENT

This research was conducted in the Future Industrial Services (FutIS) research program, managed by the Finnish Metals and Engineering Competence Cluster (FIMECC), and funded by the Finnish Funding Agency for Technology and Innovation (TEKES), research institutes.

REFERENCES

Atzori, L., Iera, A., and Morabito, G. (2010). The internet of things: A survey. *Computer networks*, 54(15):2787–2805.

Batini, C., Cappiello, C., Francalanci, C., and Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM Computing Surveys (CSUR)*, 41(3):16.

Berndt, D. J., Fisher, J. W., Hevner, A. R., and Studnicki, J. (2001). Healthcare data warehousing and quality assurance. *Computer*, 34(12):56–65.

Buda, A., Kubler, S., Borgman, J., Främling, K., Madhikermi, M., and Mirzaeifar, S. (2015). Data supply chain in industrial internet. In *Proceedings of the 11th IEEE World Conference on Factory Communication Systems*.

Dunn, S. (1998). Reinventing the maintenance process: towards zero downtime. In *Queensland Maintenance Conference Proceedings, Queensland, Australia*.

Figueira, J., Greco, S., and Ehrgott, M. (2005). *Multiple criteria decision analysis: state of the art surveys*, volume 78. Springer Science & Business Media.

Främling, K. (1996). *Modélisation et apprentissage des préférences par réseaux de neurones pour l'aide à la décision multicritère*. PhD thesis, INSA Lyon.

Främling, K., Holmström, J., Loukkola, J., Nyman, J., and Kaustell, A. (2013). Sustainable plm through intelligent products. *Engineering Applications of Artificial Intelligence*, 26(2):789–799.

Jarke, M. and Vassiliou, Y. (1997). Data warehouse quality: A review of the dwq project. In *IQ*, pages 299–313.

Kahn, B. K., Strong, D. M., and Wang, R. Y. (2002). Information quality benchmarks: product and service performance. *Communications of the ACM*, 45(4):184–192.

Knight, S.-A. and Burn, J. M. (2005). Developing a framework for assessing information quality on the world wide web. *Informing Science: International Journal of an Emerging Transdiscipline*, 8(5):159–172.

Krogstie, J., Lindland, O. I., and Sindre, G. (1995). Defining quality aspects for conceptual models. *Proceedings of the IFIP8.1 Working Conference on Information Systems Concepts: Towards a Consolidation of Views*, 1995:216–231.

Lindland, O. I., Sindre, G., and Solvberg, A. (1994). Understanding quality in conceptual modeling. *Software, IEEE*, 11(2):42–49.

Mardani, A., Jusoh, A., and Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications–two decades review from 1994 to 2014. *Expert Systems with Applications*.

Mobley, R. K. (2002). *An introduction to predictive maintenance*. Butterworth-Heinemann.

Peabody, J. W., Luck, J., Jain, S., Bertenthal, D., and Glassman, P. (2004). Assessing the accuracy of administrative data in health information systems. *Medical care*, 42(11):1066–1072.

Price, R. J. and Shanks, G. (2009). A semiotic information quality framework: Theoretical and empirical development.

Saaty, T. L. (1980). The analytic hierarchy process: planning, priority setting, resources allocation. *New York: McGraw*.

Saaty, T. L. (1990). How to make a decision: the analytic hierarchy process. *European journal of operational research*, 48(1):9–26.

Saaty, T. L. (1996). *Decision making with dependence and feedback: The analytic network process*, volume 4922. RWS publications Pittsburgh.

Tzeng, G.-H. and Huang, J.-J. (2011). *Multiple attribute decision making: methods and applications*. CRC Press.

Wang, R. Y. and Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of management information systems*, pages 5–33.

Zheng, G., Zhu, N., Tian, Z., Chen, Y., and Sun, B. (2012). Application of a trapezoidal fuzzy ahp method for work safety evaluation and early warning rating of hot and humid environments. *Safety Science*, 50(2):228–239.