eMaintenance Ontologies and Data Production

Mustafa Aljumaili
Luleå University of Technology
SE-971 87 Luleå, Sweden
+46 (0)920 49 15 53
Mustafa.Aljumaili@ltu.se

Karina Wandt
Luleå University of Technology
SE-971 87 Luleå, Sweden
+46 (0)920 49 24 83
Karina.Wandt@ltu.se

Ramin Karim
Luleå University of Technology
SE-971 87 Luleå, Sweden
+46 (0)920 49 23 44
Ramin.Karim@ltu.se

ABSTRACT
The demands on the integrity and interoperability of data in maintenance are constantly increasing. For many enterprises, achieving these demands in their operation and maintenance processes is challenging. In order to enable information exchange between systems and their actors, using effectively and efficiently-usable way, there are two essential issues which need to be addressed: i) content structure - which addresses with the description of the content's inherent elements and also the relationship between these elements; and ii) communication interface – which addresses the interface through which each specific content will be exposed. There are a lot of eMaintenance ontologies which may help to solve these issues.

The aim of this paper is to: i) explore the main ontologies related to eMaintenance solutions and to inspect how Data Quality (DQ) aspect is considered within these ontologies; ii) propose a process called "Maintenance Data Production" and to find the relation between ontologies and their role during data production stages.

Keywords
eMaintenance, Ontology, Standards, Data Quality (DQ), ICT, Interoperability, Data Production**.**

1. INTRODUCTION
Interoperability is one of the most important aspects inside companies and business enterprises. It can be defined as the ability of applications and systems to share information and exchange services with each other based on standards and to cooperate in processes using the information and services [1]. IEEE has defined interoperability as the ability of two or more systems or components to exchange information and to use that information that has been exchanged [2].

Interoperability has many objectives. One important objective is the vision of software components working smoothly together, without regard to details of any component's location, operating system, programming language, or network hardware and software [3]. One solution to achieve interoperability is using standards. As essential property for long-term data retention, they offer stability in the way information is represented, and this retention issue is increasingly recognized as a costly and critical problem for industries with long product life cycles, such as aerospace [4].

As a result to the development in Information and Communication Technology (ICT), standards are required to ensure performance, conformity, and safety of new products and processes. Standards are documented agreements containing technical guidelines to ensure that materials, products, processes, representations, and services are fit for their purpose. In case of data interoperability, standards are mainly described in the form of some data fixed formats are specified [5].

Standards, data exchange models and communication protocols are important aspects in order to achieve data interoperability between different systems in maintenance, operations and inside an organization hierarchy. When developing eMaintenance solutions as support to maintenance decision-making, integration architecture for data exchange between different data sources is important [6]. The design of integration architecture is highly depended on the mechanism that defines the structure of the data elements and also describes the relation between these elements, i.e. ontology. However, ontologies have a high impact on the integration architecture of eMaintenance solutions and affect its efficiency. The purpose of this report aims to investigate the state-of-the-art in ontologies related to maintenance.

Hence, to be interoperable, components and systems must correctly interpret words used as labels and data in an appropriate context. Today we are still far from achieving the essential levels of interoperability among manufacturing system components that provides significant improvements in manufacturing efficiency [7].

The purpose of this paper is to explore ontologies related to eMaintenance. The paper describes the focus domain for the investigated ontologies. Furthermore, the paper will study the data quality aspect within these ontologies and how that aspect is ensured during data exchange between systems.

2. MAINTENANCE DATA FLOW
Information and data quality as a competitive factor in Business Networking relates to the question of the extent to which information and data quality is decisive for the success of Business Networking efforts and therefore has a value in its own right [8].

To perform prognostic or diagnostic maintenance on a specific item, eMaintenance solutions require access to a number of different data sources, including maintenance data, product data, operation data, etc. As these sources of data often operate in a heterogeneous environment, integration between the systems is problematic [9]. As illustrated in Figure 1, different types of data are collected from heterogeneous sources, such as computer maintenance management systems (CMMS) and product data management system (PDM). The data are processed and integrated through data fusion and transformed into eMaintenance information. Therefore, the quality of the data, or Data Quality (DQ), needs to be considered during the data collection process to the data fusion and integration process.

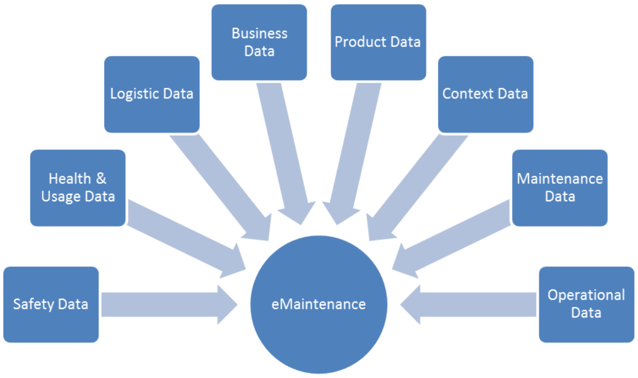

Figure 1, eMaintenance data access (Based on [9]**)**

Since data often operate in heterogeneous environments, an important aspect for eMaintenance data is interconnectivity. All systems within the eMaintenance network must be able to interact as seamlessly as possible to exchange information in an efficient and usable way. As important aspects of DQ, Data accuracy, consistency and integrity should be assured within this network [10]. Accuracy of the data means that the recorded value is in conformity with the actual value [11]. Consistency can be defined as the representation of the data value should be the same in all cases [12]. Since issues in DQ in maintenance have direct impact to decision making process, all these aspects should be considered during the data flow process. This interconnectivity issues can be solved by applying eMaintenance ontologies.

Therefore, this paper will describe the different ontologies related to eMaintenance solutions and how each ontology can assure DQ.

3. EMAINTENANCE AND DATA QUALITY

eMaintenance is a multidisciplinary domain based on maintenance and ICT, ensuring that the maintenance services are aligned with the needs and business objectives of both customers and suppliers during the whole product lifecycle [13]. eMaintenance can also be defined as a new concept that can be considered as a part of maintenance support. In this definition, eMaintenance provides information resources and information services which can be used to enable development and establishment of a proactive decision making process thought enhanced use of ICT [14]. eMaintenance solutions may include different services aimed for e.g. eMonitoring, eDiagnosis, and ePrognosis [15]. From a generic perspective, eMaintenance is maintenance managed and performed via computing [6].

Data Quality (DQ) can be defined as data that is fit for use by data consumers [16]. Information processing people considers that DQ is mainly focused on some attributes like accuracy, precision, and timeliness [17]. In maintenance decision-making, DQ is an important aspect to consider, since without control of the DQ, there is no control of the accuracy of the output [18]. In a Computerized Maintenance Management System (CMMS), three important parties may affect DQ; *data producers, data custodians,* and *data consumers. Data producers* are people or systems that generate data. *Data custodians* are people who provide and manage computing resources for storing and processing data. Finally, *data consumers* are people or systems that use data. Therefore, data users are critical in defining data quality [19]. Thus, high information content (i.e., accurate, complete, and relevant information) leads to better product cost control and increased organizational efficiency (i.e., increased profit margin, increased decision making efficiency) [20]. As mentioned before, when developing eMaintenance solutions as support to maintenance decision-making, integration architecture for data exchange between different data sources is important [6]. Therefore, eMaintenance ontologies, like standards, have contributed in enhancing maintenance DQ. In this paper, the DQ aspect within these ontologies will be investigated in an attempt to identify the main contribution of these ontologies to the DQ.

4. DATA QUALITY ISSUES

In literature, DQ has been investigated extensively in prior information science research, where much of the discussion has been devoted to the underlying dimensions, such as accuracy, completeness, presentation, and objectivity where the main focus on describing DQ these dimensions or attributes [21].

Strong et al (1997) have observed 10 key DQ problems. Their observations were based on research that employed qualitative data collection and analysis techniques. In summary, these problems are [22]:

1) Multiple sources of the same information produce different values.
2) Information is produced using subjective judgments, leading to bias.
3) Systemic errors in information production lead to lost information.
4) Large volumes of stored information make it difficult to access information in a reasonable time
5) Distributed heterogeneous systems lead to inconsistent definitions, formats, and values.
6) Nonnumeric information is difficult to index
7) Automated content analysis across information collections is not yet available.
8) As information consumers' tasks and the organizational environment change, the information that is relevant and useful changes
9) Easy access to information may conflict with requirements for security, privacy, and confidentiality
10) Lack of sufficient computing resources limits access

In general, two types of information sources can be considered: subjective and objective. Subjective sources such as human observers, intelligence agents, newspaper reporters, experts and decision makers, supply observations, subjective beliefs, hypotheses, and opinions about what they see or learn. Quality of objective information sources such as sensors, models, automated processes is free from biases inherent to human judgment and depends only on how well sensors are calibrated and how adequate models are [23].

In our previous study and from maintenance point of view, we have found that DQ issues can also be divided into two types: human related (subjective) and machine related (objective). Most of the identified DQ issues are subjective. For more information, see [24].

5. METHODOLOGY

The motivation to this study was the observed DQ issues from our previous research that has been mentioned before in section (4). In order to accomplish this study, we have divided the work into three parts. In part one, a literature study has been done to explore the eMaintenance ontologies in the different areas like condition based monitoring, product management, maintenance data management, data exchange etc. This study is done in section 4.1 eMaintenance ontologies. The second part of this study is to propose the maintenance data production process. In this process, data has been considered as a product. When we deal data as a product, DQ insurance will be easier to achieve. The last part of this study is to find relation between this process's stages and eMaintenance ontologies in order to produce high quality data. Part two and three is available in section 4.2

5.1 eMaintenance Ontologies

In eMaintenance solutions, the design of integration architecture mechanism defines the structure of the data elements and relation between these elements, i.e. ontology.

The term "manufacturing interoperability" refers to the ability to share technical and business information seamlessly throughout an extended manufacturing enterprise [4]. Published standards offer some stability by proposing information models for data representation, an essential property for long-term data exchange and archiving [25].

In this study, a literature review has been conducted in order to explore most of the available ontologies that are related to eMaintenance solutions. The studied ontologies are listed in Table 1 below. These ontologies are related to eMaintenance different solutions. The eMaintenance relation to these ontologies is stated in the right hand column of the table.

We can notice from Table 1 that eMaintenance ontologies have an important role to insure DQ in maintenance. They support maintenance DQ from data collection to the data visualization step. We can see the different eMaintenance areas related to these ontologies.

Table 1. Ontologies eMaintenance scope

Studied Ont.	Description	eMaintenance Scope
OPC UA	OPC is designed for Open Productivity and Connectivity in industrial automation and enterprise systems that support industry [26].	Software and Information interoperability
MIMOSA	MIMOSA provides metadata reference libraries and a series of information exchange standards using XML and SQL [27].	Measurement and condition based maintenance data transfer
PLCS	PLCS (Product Life Cycle Support) specifies an information model used for the exchange of assured product and support information throughout the entire product life cycle from concept to disposal [28].	Product Management and Maintenance, suitable to complex products and large companies
ISA-95	ISA-95 is the international standard for the integration of enterprise and control systems. ISA-95 consists of models and terminology to determine which information has to be exchanged between systems for maintenance and quality.	Maintenance Data Transfer and Management
XML	XML is a simple text-based format for representing structured information: documents, data, configuration, books, transactions, invoices, and much more [29].	Maintenance Data representation and exchange.
STEP	STEP is a family of standards defining a robust and time-tested methodology for describing product data throughout the life cycle of a product [30].	Product Life Cycle Management
CORBA	CORBA specifies interfaces that allow seamless interoperability among clients and servers under the object-oriented paradigm [31].	Services and Applications interoperability (objects)
OAGIS	OAGIS standard aims to achieve interoperability between disparate enterprise business systems by standardizing the architecture of the messages they exchange [32].	Data Interoperability between systems and databases
DPWS	DPWS is a common web services middleware and profile for devices, which defines two fundamental elements: the device and its hosted services [33].	Devices information exchange through web services
S1000D	It is an international specification for the procurement and production of technical publications. The S1000D provides ontology for the content of technical publications and also a content model, based on XML schema [34].	Technical Content Management
SOA	SOA represents a design framework for construction of information systems by combination of service. A service is a program unit which can be called by standardized procedures, and can independently execute assigned function [35].	Standardized data format for information and service interchange
SOAP	It defines a simple mechanism for expressing application semantics by providing a modular packaging model and encoding mechanisms for encoding data within modules [36].	Mechanism for exchanging structured and typed information between peers.
SCADA	It refers to a system that collects data from various sensors at factories, plants or in other remote locations and controls equipment over the SCADA networks [37].	Data collection and equipment control
ATA iSpec 2200	Is a global aviation industry standard for the content, structure, and electronic exchange of aircraft engineering and maintenance information. It consists of a suite of data specifications pertaining to maintenance requirements and procedures and aircraft configuration control [38].	Aviation industry
S4000M	S4000M is an International Specification for Scheduled Maintenance Analysis [38].	Maintenance Analysis
DAIS	Data Acquisition from Industrial Systems issued by the OMG is intended for online data transfer [2].	Data Acquisition

5.2 Maintenance Data Production

The Condition Based Monitoring (CBM) program to handle data and information in maintenance consists of three key steps; data acquisition, data processing and maintenance decision making [39]. This process in figure 2 includes data acquisition and processing before making a decision. Data acquisition is a process of collecting and storing useful data (information) from targeted physical assets for the purpose of CBM. Condition monitoring data are the measurements related to the health condition/state of the physical asset. It can be vibration data, acoustic data, oil analysis data, temperature, pressure, moisture, humidity, weather or environment data, etc. Various sensors, such as micro-sensors, ultrasonic sensors, acoustic emission sensors, etc., have been designed to collect different types of data. Wireless technologies, such as Bluetooth, have provided an alternative solution to cost-effective data communication. Maintenance information systems, such as computerized maintenance management systems (CMMS), enterprise resource planning systems, etc., have been developed for data storage and handling [39].

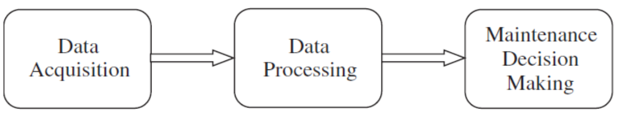

Figure 2, CBM program steps [39]**.**

However, this process can be extended to add other fundamental aspects which need to be considered when establishing a maintenance information logistics solution. These new aspects are: content, time, communication and context [14].

Based on this process, we have proposed a process called "Maintenance Data Production" to refer to the process by which data is produced. This process describes all the actions that will occur to produce the data to the consumer as a final product. Maintenance Data Production (MDP) includes the following stages: data collection, data transition, compilation, analysis, visualization and contextualization.

These stages can be summarized as follows:

1) Data collection step: when obtaining the relevant data and managing its content. This data can be collected from different sources. These sources may be sensors, RFID tags, people etc.
2) Transition step: where the collected data need to be transferred without affecting its content. Data is transferred from source location to data management system.
3) Compilation step: to compile data from different sources in a way that insures its quality.
4) Analysis step: to analyse data and extract information and knowledge for decision making support.
5) Visualisation step: to visualise the information for the intended user or decision maker. The visualization could be statistical or reports.
6) Contextualization step: to put the visualized information into the needed context so it becomes meaningful and understandable in the right context.

During our study, we tried to find the relation between eMaintenance ontologies and the maintenance data production process. The results of this study are summarized in figure 3. Figure 3 shows that eMaintenance ontologies are required in every stage of the data production. In this figure, these ontologies are located in the stage that they may affect.

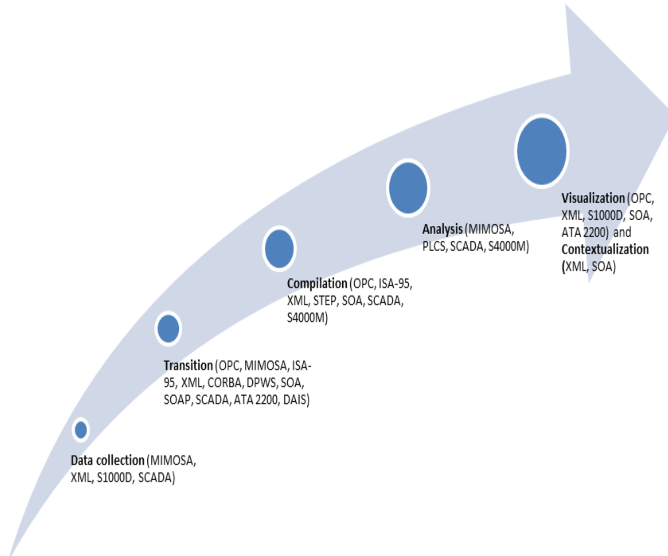

Figure 3, Maintenance Data Production Process

Hence, DQ management means that DQ should be insured during all these stages. From Figure 2, we can notice that eMaintenance ontologies have very important contribution in enhancing DQ.

6. RESULTS AND DISCUSSIONS

This study summarizes the available eMaintenance ontologies used to enhance maintenance DQ. When accomplishing the study, we can say that eMaintenance solutions are widely used and they have a lot of contributions in different industry and business aspects.

From this study, it was clear that eMaintenance ontologies contributions are represented mainly with DQ enhancement during the process of data production phases from data collection to data visualization and contextualization process. In addition, the studied ontologies were divided into the different areas it may affect.

The other contribution of this study is the suggested process that is called "Maintenance Data Production". This process considers all the stages required to produce the data as a final product. An attempt to investigate which eMaintenance ontologies are used during each stage has been done.

When applying eMaintenance ontologies to the maintenance data production process, the wide range of available ontologies makes it difficult to decide which ontology that is best suited for each stage in the Data Production process. In this article, an attempt to find the relation between the ontologies and the area that ontologies should be used in has been done. The results are shown in Table 1 above.

By applying ontologies to eMaintenance solutions according to the areas described in Table 1 and the stages in the Data Production Process, we can insure that we will get a high quality data and that leads directly to obtain an adequate and effective decision making.

7. CONCLUSIONS

Interoperability and data quality are crucial and indispensable for both maintenance and operations. They are also critical for achieving higher levels of organizational interoperability that are required for an effective decision making. In order to increase the economic benefit and enhance decision making, the use of eMaintenance tools should be adapted in more effective use. Standards, models and services discussed in this paper are offering a sustainable support for this objective. From our study, we can conclude the following:

1) DQ issues in maintenance have direct impact to the decision making that leads to wrong decisions regarding the maintenance process.
2) All discussed standards and services provide important support to the interactions necessary to construct unified maintenance and operations and enhance integration among systems of different origins to achieve integrity that helps in ensuring data quality.
3) The proposed Maintenance data production process provides a clear view to the steps that should be followed in order to produce a high quality data.
4) Applying eMaintenance ontologies in each step of the maintenance data production process helps to insure DQ.
5) Different challenges still available need to be discussed more effectively in order to help in developing adapting the available eMaintenance solutions. Some of them are represented by the high cost and the time that prevent some companies from utilizing these ontologies. Another challenge could be the lack or the ineffective use of these tools in many enterprises and organizations. Finally, the wide range of available ontologies makes it difficult to decide which ontology that is best suited for each stage in the Data Production Process.

8. REFERENCES

[1] E. Murphy, "OPC enables interoperability," *Hydrocarbon Process.*, vol. 86, no. 3, pp. 51-52 2007.

[2] OMG. *Data Acquisition from Industrial Systems* [Online]. available: www.omg.org.

[3] J. Siegel, "OMG overview: CORBA and the OMA in enterprise computing," *Commun ACM*, vol. 41, no. 10, pp. 37-43 1998.

[4] S.R. Ray and A. Jones, "Manufacturing interoperability," *J.Intell.Manuf.*, vol. 17, no. 6, pp. 681-688 2006.

[5] R.H. Allen and R.D. Sriram, "The role of standards in innovation," *Technological Forecasting and Social Change*, vol. 64, no. 2, pp. 171-181 2000.

[6] M. Kajko-Mattsson, R. Karim and A. Mirijamdotter, "Fundamentals of the eMaintenance Concept," in *1st international workshop and congress on eMainteance*, 2010, pp. 22-24.

[7] J. Klobas, "Knowledge management in small enterprises," *The Electronic Library*, vol. 18, no. 6, pp. 420-432 2000.

[8] B. Otto, Y.W. Lee and I. Caballero, "Information and data quality in business networking: A key concept for enterprises in its early stages of development," *Electronic Markets*, vol. 21, no. 2, pp. 83-97 2011.

[9] K. Wandt, R. Karim and D. Galar, "Context adapted prognostics and diagnostics," in *International Conference on Condition Monitoring and Machinery Failure Prevention Technologies*, London, UK, 2012.

[10] R. Blake and P. Mangiameli, "The effects and interactions of data quality and problem complexity on classification," *Journal of Data and Information Quality (JDIQ)*, vol. 2, no. 2, pp. 8 2011.

[11] D.P. Ballou and G.K. Tayi, "Enhancing data quality in data warehouse environments," *Commun ACM*, vol. 42, no. 1, pp. 73-78 1999.

[12] D. Ballou, S. Madnick and R. Wang, "Special section: Assuring information quality," *J.Manage.Inf.Syst.*, vol. 20, no. 3, pp. 9-11 2003.

[13] M. Kajko-Mattsson, R. Karim and A. MIRJAMDOTTER, "Essential components of e-maintenance," *International Journal of Performability Engineering*, vol. 7, no. 6, pp. 555 2011.

[14] R. Karim, *A service-oriented approach to e-maintenance of complex technical systems*, Division of Operation and Maintenance Engineering, Luleå University of Technology, 2008.

[15] U. Kumar, "Special issue on eMaintenance solutions and technologies guest editorial," *International Journal of Systems Assurance Engineering and Management*, vol. 1, no. 3, pp. 187-188 2010.

[16] D.M. Strong, Y.W. Lee and R.Y. Wang, "Data quality in context," *Commun ACM*, vol. 40, no. 5, pp. 103-110 1997.

[17] A.V. Levitin, "Data as a resource: Properties, implications, and prescriptions," *Sloan Manage.Rev.*, vol. 40, no. 1, pp. 89-101 1998.

[18] R. Karim, M. Kajko-Mattsson and P. Söderholm, "Exploiting SOA within emaintenance," in *Proceedings of the 2nd international workshop on Systems development in SOA environments*, 2008, pp. 75-80.

[19] R.L. Leitheiser, "Data quality in health care data warehouse environments," in *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on*, 2001, pp. 10 pp.

[20] N. Gorla, T.M. Somers and B. Wong, "Organizational impact of system quality, information quality, and service quality," *The Journal of Strategic Information Systems*, vol. 19, no. 3, pp. 207-228 2010.

[21] O. Arazy and R. Kopak, "On the measurability of information quality," *J.Am.Soc.Inf.Sci.Technol.*, vol. 62, no. 1, pp. 89-99 2011.

[22] D.M. Strong, Y.W. Lee and R.Y. Wang, "10 potholes in the road to information quality," *Computer*, vol. 30, no. 8, pp. 38-46 1997.

[23] G.L. Rogova and E. Bosse, "Information quality in information fusion," in *Information Fusion (FUSION), 2010 13th Conference on*, 2010, pp. 1-8.

[24] M. Aljumaili, V. Rauhala, P. Tretten and R. Karim, "Data quality in eMaintenance : a call for research," in *International Conference on Maintenance Performance Measurement & Management : MPM²*, Luleå, Sweden, 2011, pp. 69.

[25] A.F. Cutting-Decelle, L. Pouchard, B. Das, R. Young and J. Michel, "Utilising standards based approaches to information sharing and interoperability in manufacturing decision support," in *International Conference FAIM04, Toronto, Canada*, 2004.

[26] OPC. [Online]. available: www.OPCfoundation.org.

[27] MIMOSA. [Online]. available: www.mimosa.org.

[28] Eurostep. [Online]. available: www.plcs-resources.org.

[29] W3C. [Online]. available: www.w3.org.

[30] R.S. Peak, J. Lubell, V. Srinivasan and S.C. Waterbury, "STEP, XML, and UML: Complementary technologies," *Journal of Computing and Information Science in Engineering*, vol. 4, no. 4, pp. 379-390 2004.

[31] I. Pyarali and D.C. Schmidt, "An overview of the CORBA portable object adapter," *StandardView*, vol. 6, no. 1, pp. 30-43 1998.

[32] X. Fiorentini and S. Rachuri, "STEP-OAGIS harmonization joint working group," 2009.

[33] G. Cândido, F. Jammes, J.B. de Oliveira and A.W. Colombo, "SOA at device level in the industrial domain: Assessment of OPC UA and DPWS specifications," in *Industrial Informatics (INDIN), 2010 8th IEEE International Conference on*, 2010, pp. 598-603.

[34] R. Karim, O. Candell and P. Söderholm, "E-maintenance and information logistics: Aspects of content format," *Journal of Quality in Maintenance Engineering*, vol. 15, no. 3, pp. 308-324 2009.

[35] N. Komoda, "Service oriented architecture (SOA) in industrial systems," in *Industrial Informatics, 2006 IEEE International Conference on*, 2006, pp. 1-5.

[36] D. Box, D. Ehnebuske, G. Kakivaya, A. Layman, N. Mendelsohn, H.F. Nielsen, S. Thatte and D. Winer, *Simple object access protocol (SOAP) 1.1* 2000.

[37] D. Choi, H. Kim, D. Won and S. Kim, "Advanced key-management architecture for secure SCADA communications," *Power Delivery, IEEE Transactions on*, vol. 24, no. 3, pp. 1154-1163 2009.

[38] ATA. [Online]. available: www.airlines.org.

[39] A.K.S. Jardine, D. Lin and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical systems and signal processing*, vol. 20, no. 7, pp. 1483-1510 2006.