Data taxonomy to manage information and data in Maintenance Management

A. Polenghi, I. Roda, M. Macchi, A. Pozzetti

Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan 20133, Italy {adalberto.polenghi, irene.roda, marco.macchi, alessandro.pozzetti}@polimi.it

Abstract: Nowadays Maintenance Management (MM) is covering a primary role for competitiveness in manufacturing. The advent of Asset Management (AM), in which MM is a core function, enlarges the scope MM was used to. Besides, digitalization has brought a vast amount of information and data sources that MM may exploit to improve its processes and asset-related decision-making. This evolution of MM has brought a lot of opportunities but also various criticalities about information and data management. Data models are envisioned to provide significant support to this end. However, a common reference data taxonomy is needed for the correct development of data models. This work aims at exploring how the data taxonomy could help in addressing the current criticalities by synthesizing most information and data classes that support MM. The data taxonomy, along with other elements, like data models, effectively support companies in improving the management of their information and data. The usefulness of a data taxonomy is proved thanks to action research in a company within the automotive sector aiming at improving the MM process.

Keywords: information, data, taxonomy, maintenance, Asset Management, industry

1. INTRODUCTION

In the current industrial context, companies are willing to optimise at best their processes to be competitive. Among those contributing to this goal, Maintenance Management (MM) is gaining attention. An ever-increasing consciousness on the centrality of maintenance is noticed, passing from “necessary evil” to a core role in Asset Management (AM) (El-Akruti, Dwight, et al., 2013; BS EN 16646:2014, 2014), while having impact on system availability, quality, operational costs, productivity and profitability (Maletic et al., 2014). This evolution of MM over the years stems from two drivers. From one side, AM stimulates MM to consider all asset lifecycle phases (BoL – Beginning of Life, MoL – Middle of Life, EoL – End of Life) (Ouertani et al., 2008), all asset control levels (operational, tactical, strategic) (El-Akruti and Dwight, 2013), and relevant underlying principles such as life cycle, risk, system and asset-centric orientation (Roda and Macchi, 2018). From the other side, the ongoing digitalization of MM is characterised by extensive installation of sensors on assets for monitoring and controlling purposes (Bagheri et al., 2015), exploitation of advanced and sophisticated data analytics and digital twin modelling (Macchi et al., 2018), and extensive use of different IT software tools (Kans and Ingwald, 2010).

Relevant to both sides, two are the critical success factors to improve MM: human factors and information flows (Tsang, 2002). It is then essential to support the MM and AM decision-making through information (Haider, 2009), considering its flows between persons, departments and all relevant stakeholders (Polenghi et al., 2019a). However, information and data management for a proper decision support is a current challenge, in both academia and industry. To cope with this challenge, semantic data modelling is considered a promising lever (Karray et al., 2010; Negri et al., 2016). Semantic data models are ontology-related concepts that could be defined as a formal and explicit formalisation of the concepts. They are at the basis of further advances in interoperability among different enterprise systems, and even in reasoning built on inference capabilities over dispersed information and data (Fortineau et al., 2013). It is particularly relevant to set up interoperability along the asset lifecycle (Emmanouilidis et al., 2019). At the normative level, the ISO 15926 (2004) is the primary attempt to define a common data model to integrate lifecycle data in an industrial context, with a focus on process industry. This first attempt could inspire the development of a sector/industry-independent data model, taking into account other contexts within the discrete manufacturing, typically behind the average of the process industry in managing data along the lifecycle of their assets, while also using different types of assets.

The aim of the research underlying this paper is to develop an industry-independent data model. As suggested by (Kiritsis, 2013), the starting point to approach this endeavour is an underlying common data structure, or data taxonomy. In this paper, the data taxonomy is presented as a supporting tool in projects dealing with the information and data management side of a particular process; therefore, the research objective of this specific work is to show how a data taxonomy brings several benefits to the MM decision-making by means of an action research in an automotive company. The paper is so structured: section 2 lists the information and data criticalities through extant literature; section 3 describes the adopted data taxonomy; section 4 deals with the action research both as methodology and results; section 5 critically analyses the
characteristics of the automotive company that affects the MM process; finally, section 6 states some conclusions and envision future developments.

2. INFORMATION AND DATA CRITICALITIES FOR MAINTENANCE DECISION-MAKING

The number of data sources on which a company could count on is remarkably increased in the last years due to the digitalization process. Among other organisation functions, maintenance is facing a lot of opportunities in this scenario. Nevertheless, the criticalities related to information and data management are numerous (Petchrompo and Parlikad, 2019). Considering the typical steps required to this end, i.e., data collection, data to information transformation, information management and integration, we provide a summary of some criticalities according to extant scientific literature (Table 1).

Table 1. Information and data criticalities

| Step                      | Criticalities                                                                 |
|---------------------------|-------------------------------------------------------------------------------|
| Data collection           | 1DC: Heterogeneous data to be managed (different formats, different sampling frequencies, different sources, also geographically dispersed) (Mahlamäki and Nieminen, 2019) |
|                           | 2DC: No automatic transmission of data from shop-floor (also delayed) (Ćwikła et al., 2017) |
| Data to information       | 1DI: Challenging volume and variety of data to elaborate (Sharma et al., 2017) |
| transformation            | 2DI: Data quality and data compliance not respected (Kortelainen et al., 2015) |
| Information management    | 3DI: Required information and data often missing (Tretten and Karim, 2014) |
| and integration           | 1IM: Various information systems to be integrated for decision-making (Legat et al., 2014) |
|                           | 2IM: Cybersecurity to prevent uncontrolled data access and exchange (Wang et al., 2017) |

It is worth pointing out that having different data sources to manage is a major issue affecting different other steps. In the data collection step, effort is put on integrating heterogeneous data while, in the information management and integration step, additional work is also required to integrate the already elaborated information from different organisation functions, in order to support asset-related decisions.

3. ADOPTED DATA TAXONOMY

The data taxonomy aims at standardising and formalising information and data needed to MM process by defining a common structure. Besides, the taxonomy guarantees coherent development to support information and data management improvement by: i) identifying the kind of information and data needed to enhance decisions in the MM process; ii) supporting a mapping activity to identify where the raw data are stored in the company IT systems; iii) developing a data model to support the MM process and the related decisions.

The taxonomy herein used mainly derives from international standards focused on MM and AM since most of them already face information and data issues (Polenghi et al., 2019b). The ones used to build the taxonomy are ISO 5500x (2014, 2018) body of standards on AM; ISO 13306 (2017) on maintenance terminology; ISO 14224 (2016) on reliability and maintenance data exchange; ISO 15341 (2019) on maintenance key performance indicators; ISO 16602 (2014) on spare product assurance through FMEA/FMECA methodology; IEC 60812 (2018) on FMEA/FMECA methodology; IEC 61508 (2010) on electrical, electronic, programmable electronic safety-related systems. In Fig. 1, the data taxonomy is presented as a hierarchical structure starting from the top-class Asset.
The adopted data taxonomy is not constrained to consider just structured data which include mainly asset-related data such as the design specifications and asset’s attributes (manufacturer’s name, locations); instead, the requirements from stakeholders and even the registered work orders are far from being fully structured because an important part of them typically come in the form of description in natural language. Both structured and unstructured data support the entire lifecycle management of an industrial asset; then, they must be managed despite the different repositories in which they are stored or the different formats through which they are expressed. Indeed, the retrieval of all information and data is impacted by their source. In the current context, the data may be widespread in the company, stored in different repositories and formats. Some data may come directly from sensors, as variables monitoring a specific failure mode of the asset, or even from local or cloud-based databases, as work orders or events the asset underwent, or even from specific software tools, as the technical drawings of the asset/asset systems or, as other case, the procedural descriptions of maintenance actions.

The “knowledge” of all information and data classes is at the background of the project we actively do through action research aiming at increasing MM process performances. Also, the company does not only act as a receiver of the taxonomy, but it helps in improving it through the interviews.

4. ACTION RESEARCH: PROJECT SUMMARY

The action research is performed within a company in the automotive industry, selling its products for the main carmakers worldwide. The company is a world leader in designing and producing mechanical parts for braking and fuel systems. The production system of the company is characterised by two areas, machining and assembly areas, with a supermarket (interoperation buffer), as in Fig. 2. The products flow from machining area, through supermarket to assembly area (from bottom to top of Fig. 2).

The first step of the project is to map the current company MM process (subsection 4.1), then the adopted methodology and results are described (subsection 4.2), finally the benefits of adopting a data taxonomy are illustrated (subsection 4.3).

4.1 Overall map of the company MM process

The current map of the MM process in the company reports some criticalities that must be addressed. It is worth noting that these criticalities are of different nature:

- Uncomplete formalisation and integration of the MM process:
  - The MM process is not consistent between the two production system areas, i.e. machining area and assembly area (as explained later on);
  - There is a significant time delay between the moment in which the operator performs the maintenance action/s and the moment in which he (or who in charge) registers the work;
  - Even if some information/data are recognised as important, they are not considered, thus not formalised, in the process, e.g. the failure mechanism;
  - Internal stakeholders’ requirements are not formalised, so there is no formal alignment with other organisational units; as a side effect most of the attention is given to customer’s needs, neglecting the cooperation between organisational units.

- Heterogeneous IT systems:
  - The owned IT tools are both developed in-house and commercial solutions, that are not interoperable each other;
  - Each organisational department or units use its own IT tools not compatible with other ones.

- Data with challenging volume and variety:
  - Different volume, given a high volume of process data (up to 5 Gb/day) and low volume from quality data (up to 10 Mb/day), are available as raw data;
Different velocity, since they are both automatic (sampled at 1 MHz) or handed-inputted data (usually registered daily);

- Different variety, as there are both structured data (machine parameters), semi-structured (data filled in work orders), and un-structured (on-board quality issue description).

The latter two criticalities confirm the scenario depicted through the literature review (Table 1, points 2DC, 1DI, and 1IM). The company is facing different difficulties in managing all information and data for various reasons. The project aims at paving the way for further actions to improve MM, including the process itself. Subsection 4.2 describes the adopted methodology.

4.2 Project methodology and results

Addressing the above criticalities requires a structured methodology able to formalise the process, as well as the related IT tools and data. While on one side the process analysis could count on different well-structured and fit-for-purpose methodologies (Aguilar-Savén, 2004), the analysis of IT tools and needed data for the MM process is more unstructured. Therefore, we adopt a three-phase methodology that starts with the process formalisation, and then a detailed analysis of IT tools and data by means of interviews:

1. Process mapping;
2. IT tools mapping;
3. Data mapping.

1. The first phase, i.e. process mapping, is performed by applying the BPMN (Business Process Modelling and Notation) methodology. The BPMN is used to realise diagrams for each process of interest to the company, namely corrective and preventive maintenance. The use of BPMN is propaedeutic to provide an organisation-oriented view of the MM process since it allows subdividing the diagrams according to the organisational units in charge of the activities. The process mapping results show that today the MM process is performed by three different units: one unit from the quality department (in charge of the quality of the product) and two units from the technical department (in charge of the design and management of production systems).

2. Once the process is formalised, the IT tools mapping phase is performed. It is worth noting that the two production areas, i.e. machining and assembly areas, rely on different software for managing the maintenance process. The assembly area mainly relies on the in-house developed MES (Manufacturing Execution System), while the machining area is managed through the ERP (Enterprise Resource Planning) module for MM. This situation exacerbates the information and data exchange and sharing problems of the company.

3. Finally, the data mapping phase is realised: each MM process step needs either to use elaborated data either raw data. Table 2 provides an example of the result obtained by phases two and three of the methodology: the first column reports a branch of the preventive MM process for the assembly area (in accordance to current company steps’ nomenclature, TPM = Total Productive Maintenance), the second column the used IT tools, and the third column the data and information needed. Data/Info column is organised according to the classes in the data taxonomy as main driver.

### Table 2. Example of results for phases two and three

| Process steps | IT tools     | Data/Info                          |
|---------------|--------------|------------------------------------|
| Carry out TPM assessment | Microsoft Excel | Asset Data:                        |
| Schedule TMP assessment | MES          | Line section/Station                |
|                |              | Asset type                         |
|                |              | Asset subunit                      |
|                |              | Maintenance Data:                  |
|                |              | Supervisor                         |
|                |              | Operator                           |
|                |              | Needed spare part/s                |

In the specific example of Table 2, the reported process step is related to the assessment of TPM to determine the maintenance plans. For the sake of shortness, the data and information in the right-hand column are partial with respect to the ones really used by the company in the analysed phase in the provided example.

After the application of the methodology, the project has a series of outputs that could be summarised in the remainder, as validated by the company asset manager:

- better formalisation and standardisation of the MM process, which involves either a better structuring of MM steps already performed, and suggestions for improvements in some parts of the process;
- formalisation and standardisation of the information and data needed in each step of the MM process.

As a result of these two outputs, a best practice is realised that will help future application of the MM process to different assets, including new purchased assets as well as those already installed, but whose maintenance plant has not been recently revised.

4.3 Benefits of adopting the data taxonomy

The adoption of a data taxonomy helps in addressing most of the phases of the methodology, especially IT tools mapping and data mapping, namely phases two and three, respectively. During the project, different are the benefits recognised by the company experts in this regard. The data taxonomy serves either to guide the phases and to state some intermediate results as well. As a guideline, it helps managers, operators and researchers in:

1. getting the wide map of all information and data needed for a suitable and comprehensive MM; from the very beginning everyone in the project was aware of the entire set of information and data needed, from traditional maintenance-related ones, as failure data, to the new ones inspired by AM, as business context and stakeholders’ requirements;
2. retrieving all information and data from different sources, avoiding missing something; it relates also to point 1 since the clear map of information and
data speeds up listing all IT tools, databases, and repositories in general, interesting for MM process.

As a “tool” for intermediate results, it supports in:

1. understanding the current situation in terms of data completeness in the MM process; for example, the company was aware of most of the information and data needed for MM, but with the complete map provided by the taxonomy, they realise that all failure characteristics like failure mechanism are neither formally registered nor considered;
2. enriching the current MM process with information and data not considered so far; for instance, the importance of registering all alerts (not yet registered) and alarms to be then analysed is made evident by the taxonomy.

These intermediate results are extremely useful to the company which realises the need to put resources to improve the current policies of information and data management, even beyond the MM process.

5. CRITICAL DISCUSSION

Even though the project is successful, it reveals some intrinsic criticalities the company is facing, mainly relating to the industry in which it is acting and its historical evolution.

First and foremost, the quality that must be guaranteed to carmakers is so high that the entire supply chain is controlled. This circumstance reflects in the internal processes of the company since most of them are followers of the quality department that, in the end, guides the entire company strategy. However, this internal unbalancing towards a quality-driven MM process shapes the maintenance function as a non-value-added activity that must always chase the customer’s claims in a reactive way rather than proactively.

Secondly, all the information and data nowadays collected by the companies are structured for root cause analysis to answer customer’s claims. This prevents maintenance to easily use those data to improve its performance since most of them are product-related rather than process-related.

Finally, the high customisation of the final product creates a technological divide between the machining and assembly areas. In the first one, the machines are old (up to 25 years old) and not completely connected; this, among other factors, prevents maintenance from acting on machines’ failures proactively. In the last one, the machines are designed according to each customer and to its specific product model; it results that the machines are quickly replaced and so they are up to date with newly available technologies.

6. CONCLUSIONS

Among several criticalities related to information and data management, the integration of different sources is seen as a current lack to overcome and so suitably support the MM decision-making. To this end, data models offer a promising way to fulfil this challenge, but they must be supported by a proper data structure, or taxonomy, helping in standardising information and data. The adopted data taxonomy aims at supporting data modelling but, before being instantiated in a data model, it offers by itself several positive outputs. From one side, it guides company experts and researchers involved in the project to get the wide map of all the information and data needed and the relative sources. From the other side, it provides the state of practice of a company about information and data completeness with respect to the MM process. Thus, suggesting additional data to be gathered goes in the direction of improving the whole information and data management strategy of the company, with the final aim of enhancing and boosting asset-related decisions.

At the end of the project shown herein as action research, the MM process is more standardised and formalised, as well as its related information and data. The implications it has are mainly related to the collection of data and improvement of systems (including IT tools and databases) interoperability, where needed. For instance, the company already plans a path to improve MES to collect feedback from operators on the failure in a more structured way (e.g. by already defining and listing the failure modes rather than leaving it to the uncontrollable and error-prone interpretation of the operator).

The data taxonomy is demonstrated to be a useful and practical tool either for researchers and companies interested in MM and AM. It aims to be general, even if it is tested in a project confined in the automotive industry. Anyhow, it must be remarked that this represents only a first step towards the integration of information and data sources. Once the data structure will be consolidated, and coherency and consistency of the data classes verified against different contexts/sectors, an integrated data model could be drawn. It may well represent the MM process in all its facets, and it could support overcoming also other information and data management criticalities nowadays hugely affecting different companies. As a long-term vision, ontologies applied to the MM field may definitely guide the integration of different information and data sources as well as properly supporting the decision-making process.

REFERENCES

Aguilar-Savén, R. S. (2004) ‘Business process modelling: Review and framework’, *International Journal of Production Economics*, 90(2), pp. 129–149. doi: 10.1016/S0925-5273(03)00102-6.

Bagheri, B., Yang, S., Kao, H.-A. and Lee, J. (2015) ‘Cyber-physical Systems Architecture for Self-Aware Machines in Industry 4.0 Environment’, *IFAC-PapersOnLine*. Elsevier, 48(3), pp. 1622–1627. doi: 10.1016/J.IFACOL.2015.06.318.

BS EN 16646:2014 (2014) ‘Maintenance — Maintenance within physical asset management’, *BSI Standards Publication*.

Ćwikla, G., Gwiazda, A., Banaś, W., Monica, Z. and Foit, K. (2017) ‘Analysis of the possibility of SysML and BPMN application in formal data acquisition system description’, in *IOP Conference Series: Materials*
