CIGI QUALITA MOSIM 2023

Data quality challenges in dashboard development in industry 4.0 context: an engineering-to-order manufacturing case study

Difanrsé Rebecca HOUITTE[1], Tasseda BOUKHERROUB[1,2], Darine AMEYED[3], Alexandre MOISE[4]

[1] DEPARTMENT OF SYSTEMS ENGINEERING & NUMERIX LABORATORY
École de Technologie Supérieure (ÉTS), Canada
École Centrale Méditerranée, France
difanrse-rebecca.houitte.1@ens.etsmtl.ca

[2] INTERUNIVERSITY CENTER ON ENTREPRISE NETWORKS, LOGISTICS AND TRANSPORTATION (CIIRELT)
Université de Montréal, Canada
tasseda.boukherroub@etsmtl.ca

[3] SYSTEMS ENGINEERING DEPARTMENT & SYNCHROMEDIA LABORATORY
École de Technologie Supérieure (ÉTS), Canada
Université du Québec, Canada
darine.ameyed.1@ens.etsmtl.ca

[4] BUSINESS SCHOOL, CREATECH RESEARCH CENTER
Université de Sherbrooke, Canada
alexandre.moise@usherbrooke.ca

Abstract – The use of dashboards has become increasingly prevalent as a means of providing real-time information and decision-making support across various industries, particularly in the context of industry 4.0. This study examines the feasibility of creating a dashboard for an engineering-to-order (ETO) manufacturing factory to support its transition toward a smart factory. It highlights the challenges that could be encountered in this process. In ETO factories, products are highly customized to match the specifications of the clients, which can make performance monitoring difficult. Our study provides a roadmap for manufacturing factories aiming to develop their dashboard and transform themselves into smart factories. It shows that a rigorous analysis of data quality is essential to ensure the accuracy and reliability of the data required for performance measurement before creating a dashboard.

Keywords – Dashboard, data quality, engineering-to-order, smart factory, performance indicator.

Résumé – L'utilisation de tableaux de bord est devenue de plus en plus courante dans diverses industries, notamment dans le contexte de l'industrie 4.0, comme moyen de fournir des renseignements en temps réel et comme soutien à la prise de décision. Cette étude examine la faisabilité de créer un tableau de bord pour une usine de fabrication qui fonctionne en mode conception sur commande (ETO) afin de soutenir sa transition vers une usine intelligente. Elle met en lumière les défis qui pourraient être rencontrés dans ce processus. Dans les usines ETO, les produits fabriqués sont hautement personnalisés selon les spécifications des clients, ce qui peut rendre difficile le suivi de la performance. Notre étude fournit une feuille de route pour d'autres usines de fabrication souhaitant développer leur tableau de bord et se transformer en usines intelligentes. Elle montre qu'une analyse rigoureuse de la qualité des données est une étape essentielle avant de créer un tableau de bord, afin d'assurer l'exactitude et la fiabilité des données requises pour la mesure de la performance.

Mots clés – Tableau de bord, qualité des données, conception sur commande, usine intelligente, indicateur de performance.

1 INTRODUCTION

In recent years, the industry has undergone a significant transformation with the emergence of advanced technologies such as Artificial Intelligence (AI), Internet of Things (IoT), and robotics, leading to what is known as Industry 4.0 or digital transformation [Rochas et al., 2022]. This transition offers new opportunities for data-driven decision-making and the development of tools to ease decision-making. An example of these tools is the dashboard. A crucial aspect of creating a dashboard is ensuring that the data being collected to perform visualizations is of good quality. A dashboard that is based on inaccurate or unreliable data will provide a false or misleading representation of the manufacturing process, which can lead to poor-quality decisions and inefficiencies. However, obtaining good quality data can be challenging as it often comes from multiple sources and may contain errors, inconsistencies, spelling mistakes, or missing information. In this article, data quality issues refer not only to these aspects, but also to the lack of data.

Our study aims to analyse the feasibility of creating a dashboard for an engineering-to-order (ETO) manufacturing factory in order to support its transition toward a smart factory. Specifically, we examined the challenges encountered in the process of creating a dashboard for operational and tactical use for the ÉTS (École de technologie supérieure) Institutional Manufacturing Laboratory (LIFE), located in Montréal (province of Québec), Canada, and proposed solutions and recommendations to address these challenges. Two major challenges were identified: the lack of data, and the poor quality of the data collected. Our methodology and recommendations could serve as a roadmap for other manufacturing factories aspiring to develop their dashboards and evolve into smart factories. Indeed, the scientific literature lacks studies discussing the process of creating a dashboard, notably for ETO factories and challenges encountered during this process such as those presented in this study. Therefore, our study contributes to filling this gap.

The remainder of this article is organized as follows: we review the literature related to dashboard development and data quality issues in section 2. Section 3 describes our methodology. Section 4 presents the performance indicators selected for LIFE factory as well our solutions and recommendations to address some of the data quality issues identified. Section 5 concludes the paper and adduces some research perspectives.

2 LITERATURE REVIEW

The dashboard is defined as “a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance” [Few, 2007]. In the manufacturing context, a dashboard can present a collection of visualization items such as graphs, charts, and tables, that provides an overview of measures and key performance indicators (KPIs) relevant to the manufacturing process and to decision-making. A KPI is a performance measurement that is in line with the business strategy and has clearly defined targets, time frames, and baselines. What sets it apart from regular metrics is that it reflects a strategic objective and assesses performance against a specific goal [Eckerson, 2011]. Dashboards can be classified into three categories according to the level of decision support: strategic, tactical and operational. These categories are not mutually exclusive [Sarikaya et al., 2019]. Strategic dashboards provide a long-term view and are used to track KPIs for the entire organization. Tactical dashboards, on the other hand, provide a mid-term perspective and are used to track progress towards specific goals. Finally, operational dashboards focus on the day-to-day monitoring of a specific process or activity and are used by front-line workers [Rahman and al., 2017].

The use of dashboards has become increasingly prevalent as a means of providing real-time insights and decision-making support across various industries, notably in the context of intelligent factories. However, designing and implementing dashboards can be challenging. It is important to consider the specific needs and goals of the organization when designing them [Rahman et al., 2017]. For [Larsson et al., 2021], after identifying the goals, the next step to create a dashboard is conducting a comprehensive literature review of performance indicators, particularly related to the organization's work environment, and then selecting the relevant ones. These indicators are determined by the type of metrics that decision-makers need to access in order to make decisions [Tokola et al, 2016]. A survey can be a useful tool for identifying the most important performance indicators for a manufacturing organization. For example, a survey on 99 indicators performed by [Tokola et al., 2016] shows that the most important ones for manufacturing organizations are related to efficiency, quality, and production costs. Once these metrics have been identified, the required data to measure them must be collected.

In the context of Industry 4.0, as new technologies are integrated into the manufacturing processes, it is common to use pre-existing data and add data gathered from additional new sources. According to [Cerquitelli et al., 2021], data acquisition should consider three sources: 1) data that already exists in the manufacturing company, 2) data that is generated directly by the manufacturing machines/equipment, and 3) data collected through sensors that are added to machines specifically for this purpose. For [Williams and Tang, 2020], the greatest challenge is ensuring the integrity of the data. They point out that the big data produced in manufacturing environments are heterogeneous and complex, and errors or inconsistencies can impact the quality of insights generated from that data. The data quality dimensions are characteristics of data that reflect its suitability for use. By evaluating data against these dimensions, organizations can determine the degree to which their data meets the requirements for a specific purpose or use [Cichy and Rass, 2019]. They help ensure that data is reliable, accurate, and fit for its intended use, which is essential for making informed decisions and driving business performance.

According to Cichy and Rass's survey [Cichy and Rass, 2019], the most commonly used dimensions in data quality frameworks are completeness, timeliness, accuracy, consistency, and accessibility. That [Ly, 2020] defines them as follows. Completeness refers to the extent to which all appropriate data is available for analysis. Timeliness refers to how quickly data is acquired and used after it is generated, as well as how up-to-date it is. Accuracy refers to the degree to which the data reflects the true state of the system or process being measured. Consistency refers to the degree to which data is presented in the same format and has the same meaning even when integrated across different systems and sources. Accessibility is “the extent to which information is available, or easily and quickly retrievable”.

Our literature review highlights the important role of high-quality data in the development of effective dashboards in manufacturing contexts. However, we could not find specific information on the design of dashboards for ETO factories. Indeed, we find just one article discussing the process of making a balance scorecard and a digital dashboard using an ETO factory as a case study [Larsson et al., 2021]. But it did not provide details on the unique characteristics of the ETO environment.

3 METHODOLOGY

Our study followed a step-by-step method (as depicted in Figure 1) which can serve as a framework to guide other manufacturing factories seeking to develop their own dashboard.

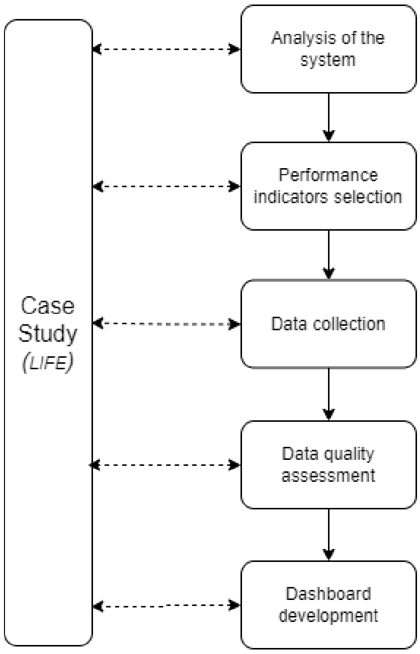

Figure 1. Methodology to create a manufacturing dashboard

3.1 Overview of the proposed method

- Analysis of the system

The initial step in our project is to analyse the system comprehensively to gain an understanding of the factory's mission, operation process, and communication system. Firstly, understanding the factory's mission and goals is essential because any project must align with the overall objectives of the factory to be successful. Secondly, it is important to understand the operation process of the factory to create a useful dashboard. Indeed, the dashboard must be tailored to the specific needs of the factory. Finally, the dashboard will be part of the communication system and must integrate seamlessly with existing communication channels. That is why we must also focus on understanding the communication system and information flow of the factory. In our case study, we achieved this goal by interviewing the LIFE management team. The process resulted in the creation of detailed manufacturing process maps that will also serve for future projects.

- Performance indicator selection

This step consists in identifying performance indicators relevant to the factory context. It involves a comprehensive review of the scientific literature, to gather the potential indicators for the factory. Then the relevant indicators can be selected based on information gathered during the analysis of the system (first step). Our selection implied a survey and an interview conducted with the management team and employees. This approach ensured that we considered a broad range of perspectives and sources of information, resulting in the identification of meaningful performance indicators.

- Data collection

The next step is to collect the data needed to "feed" the chosen indicators. It involves analysing all available data in the factory and determining the data necessary to measure the indicators. This can be a challenging task due to the volume of data available, and not all data may be relevant or useful in measuring the identified indicators. In some cases, we identified that new data acquisition methods such as real time data sensing can be very useful to gather additional information. However, this is beyond the scope of our study.

- Data quality assessment

The data quality assessment is important as data quality issues directly impact the value of the insights and decisions to be made. Based on [Cichy and Rass, 2019], we selected and used five data quality dimensions to assess our data: completeness, timeliness, accuracy, consistency, and accessibility.

- Dashboard development

The final step is the development of the dashboard. Although this step is the ultimate goal of this research project, it is not the focus of this paper.

3.2 Case study

In our project, we used the LIFE factory as a case study. LIFE is an ETO manufacturing factory specialized in modelling, designing, and manufacturing customized products to meet specific customer requirements (ÉTS community such as researchers, teachers, and students). Over the past 18 years, the factory has supported various teaching and research projects for the ÉTS community, delivering more than a hundred projects annually. It has a team of 15 employees, with a diversity of advanced machinery and cutting-edge technologies, including digital controlled machines, 3D printers, robotic machining cells, etc. The need for better material and data flow management, communication, planning, decision support, and prediction tools is high at LIFE. Developing an effective dashboard aligns with the factory's goals of remaining at the forefront of the smart/connected factory movement and Industry 4.0 transition.

3.3 Analysis of the system

To gather information about the factory, we visited it and interviewed the management and executive team. We identified the main goals of the LIFE factory and the roles of all the employees. We then focused on the operation process. To describe it clearly, we created two distinct maps (on Microsoft software Visio), a Swimlane detailed process diagram, and a BPMN (Business Process Model and Notation) diagram [Object Management Group, 2011], in collaboration with one of the managers who reviewed and validated them over time. The Swimlane diagram is useful for showing the way the factory works. The BPMN diagram helps to know where the information can be collected throughout the process and to identify the communication channels. By creating these maps, we sought to make the factory operation process more transparent and accessible to everyone.

3.3.1 Mission of the factory

As an ETO factory, the main goals of LIFE are to improve product delivery delays and ensure high product quality. According to [Hicks et al., 2000], to improve delivery performance, the ETO factory must focus on two factors: shortening the time from order to delivery (lead time) and increasing the precision of lead-time predictions. After identifying the missions of the factory, we focus on the roles of all stakeholders involved, including the clients, managers, shop-floor employees, the executive (one person), finances, and suppliers. The four managers of the factory, which consist of one engineer/lead manager and three technicians, are responsible for direct customer contact, planning/supervising the production process, and ensuring that workers have adequate resources to complete their tasks. In addition, the lead manager oversees the overall functioning of the factory. The nine shop-floor workers are responsible for programming and setting-up the machines, manufacturing different items, and the final assembly of the items. They are directly involved in the manufacturing process. The executive is responsible for tracking annual performance and overseeing the financing aspect of the factory, including grant applications, preparing annual balance sheets, and more.

Based on the organizational structure, the management processes of the factory, which is led only by one executive, supported by the autonomy of the management team with the lead manager supervision, also the size of the factory (i.e., 15 employees), etc., and additionally our discussions with the executive and management team, we identified that a strategic dashboard should not be a priority. Therefore, we have determined that developing tactical and operational dashboards would be prioritized.

3.3.2 Operation process mapping

The Swimlane detailed process diagram is a visual representation of the detailed manufacturing process that shows the tasks and responsibilities within an organization. The chart is divided into horizontal lanes, each representing a different function within the organization. LIFE's Swimlane detailed process diagram encompasses eight main steps (macro-processes), each detailing a part of the entire manufacturing process: 1) preliminary project analysis, 2) technical analysis, 3) financial analysis, 4) production planning, 5) production, 6) assembly, 7) project delivery and 8) project monitoring. We created a simplified version of the detailed map to show the main steps of the process and the main actors involved in each step (see Figure 2).

A new project (order) begins with a meeting or exchange of messages (emails) between the manager and the client. The client provides the manager with information about the desired product, including its purpose, dimensions, expected performance. In turn, the manager provides the client with technical information about the feasibility of the project, including the estimated cost and lead time. This step is important in an ETO factory since the product quality is based mainly on the client's specification satisfaction.

The manager then works on SolidWorks to create 3D and 2D models of the items to be manufactured (in some cases, 3D or 2D models are provided by the client). The models are presented to the clients. Once they are validated, the manager makes quotes and sends them to the finances department (ÉTS). After that, he assigns the tasks required for the project to the shop-floor employees.

In the LIFE factory, a project is composed of tasks and/or external manufacturing orders. A task is a product that is manufactured in the factory and an external order is a product manufactured by an external factory. LIFE Managers use a production management software called "GestLIFE" to monitor the projects' progress. They also use the software to share relevant information with the workers on the production floor. For instance, an operator can identify in the management software the tasks assigned to him, the machines and materials involved, and where to find the 3D models.

The software provides a wealth of information, but it can be difficult to understand, and some information is not accessible. Indeed, the software is not very user-friendly for real-time monitoring, and it does not provide a clear or visual representation of the project's progress. This can make it difficult for the managers to quickly understand the status of a project or identify any issues or delays that need to be addressed. To overcome these limitations, the managers often go directly to the factory to observe the production process firsthand. This approach is both time-consuming and can be prone to human errors. A dashboard, especially a tactical one, can be used to address this problem by providing a more effective visual and accessible way to understand, for example, the progress of a given project.

The Swimlane diagram does not provide enough details about the way information is shared between the different actors and between stages. It was therefore necessary to also create a separate map for representing the information flows (BPMN) (see sub-section 3.3.3).

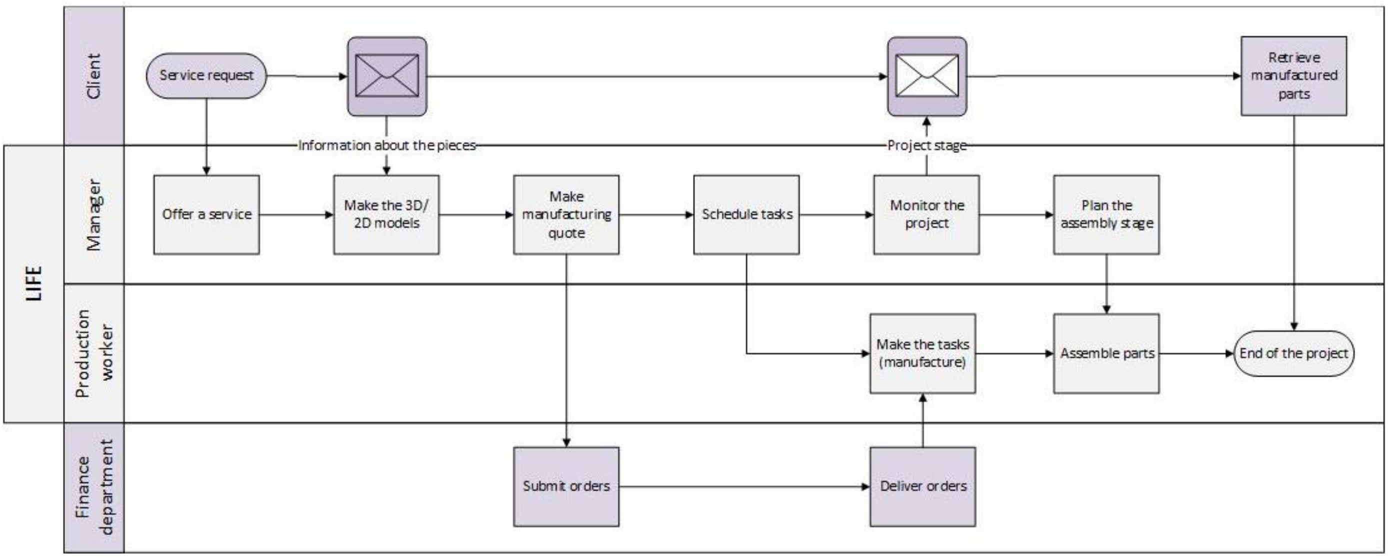

Figure 2. Simplified operation process map

3.3.3 Information flows mapping

One effective way to map the information flows is using a BPMN diagram. It allows for a clear visualization of the data transmission between different actors and systems, and for the identification of inefficiencies in the communication system. We created a BPMN diagram for LIFE factory, but it is not shown in this article due to its large size.

Mainly, the diagram shows that all customer-related data such as emails, pictures, and sketches, is organized in a dedicated folder, which contains all the artifacts resulting from the exchange of information between the customer and the manager. The production-related data is available on the management software GestLIFE that also serves as the main communication channel between managers and operators. However, the last stages of production, specifically assembly, heavily relies on verbal communication and is not adequately captured in the management software. It can be challenging to track real-time production status, which can have an impact on delivery date reliability as managers do not have enough data to make accurate predictions. This highlights the need for a more comprehensive and efficient method of information tracking and recording, particularly for management use.

3.4 Performance indicators selection

Our performance indicator selection process started with a literature review to identify indicators relevant to manufacturing facilities. Based on the results of this literature review and our understanding of the operation and needs of the LIFE factory, we proposed a set of performance indicators for tracking various aspects of the factory's performance, such as productivity, lead time, planning, and quality. The literature provided us with a solid foundation of knowledge on industrial best practices and appropriate performance indicators, while our knowledge of the factory's operations helped us to tailor our selection to the specific requirements of the factory.

We identified 18 tactical indicators (aimed for a tactical dashboard) and 12 operational performance indicators (aimed for an operational dashboard). We prepared a survey that we shared with the factory employees and managers to collect their feedback on the proposed indicators and to select the most important and useful ones. It allowed us to gather insights and perspectives from those who are most familiar with the operations process. After sending the survey three times, we received four answers, from two managers and two operators (27% of the employees). We start the survey in December 2022, which was a very busy period for the factory employees. Therefore, to obtain more feedback, we also conducted interviews with the executive who has significant experience in the factory (he was the former manager of LIFE). During the selection process, we also used the well-known SMART criteria (Specific, Measurable, Achievable, Relevant, and Time-bound), which are briefly described below:

- **Specific**: the metric should be specific to the goal or objective being measured.
- **Measurable**: the metric should be measurable, meaning that data can be collected and analyzed to track progress.
- **Achievable**: the metric should be achievable, meaning that it should be realistic to achieve with the resources and constraints of the organization.
- **Relevant**: the metric should be suitable to the goal or objective being measured.
- **Time-bound**: The metric should have a specific time frame within which progress will be measured.

By using these criteria, we aim to select performance indicators that are clear, actionable, and directly related to the goals of the organization. We realized that two SMART criteria were difficult to satisfy: some indicators were not relevant, and others were not measurable. For example, indicators related to inventory are not relevant for ETO manufacturing, because raw materials are typically ordered and used on a project-by-project basis, meaning that the exact type and quantity of materials needed can vary greatly from one project to another. This makes it difficult to establish standardized metrics for raw material inventory levels, as the inventory needs are different for each project.

Regarding the measurability, a few indicators could not be kept because there is not enough data available to track them. For example, it is not possible to measure customer satisfaction because there is no review collected from the clients. It would be important to put in place a system to collect this data in the future, especially because quality is a major performance aspect for an ETO-based factory (product quality is strongly related to customer satisfaction).

In the end, four operational and eight tactical indicators were selected (see section 4.1). However, we realized that even the data available to track these indicators present quality issues, which can lead to inaccurate or unreliable results and compromise the effectiveness of the dashboard. To address these issues, we collected the data and conducted a data quality assessment.

3.5 Data collection

As mentioned before, there are two types of data available in the factory: client-related data stored in a dedicated folder and production-related data stored in the management software GestLIFE. The client-related folder includes a variety of documents (emails, drawings, sketches, etc.), but due to the complexity and specificity of each file, we found that this data was not very useful. Instead, we focused on the data available in GestLIFE, which consists of 40 different files exported as Excel files. This data was much more structured and organized, and therefore more suitable for our analysis. We still had to navigate through the data to identify the information relevant to the selected performance indicators. Once we have identified the useful data, we performed the data quality assessment described in the next sub-section.

3.6 Data quality assessment

We selected the most common dimensions of data quality (completeness, timeliness, accuracy, consistency, and accessibility) to assess our data based on the literature [Cichy and Rass, 2019]. These are briefly described in the following:

- **Completeness**

It refers to the extent to which all data elements are collected. It can be measured by the percentage of collected elements compared to the expected number. For data to be considered complete, this percentage must reach 100%. Any missing data can lead to incomplete or inaccurate representation of the manufacturing process, which can impact effective decision-making. Specifically, in an ETO factory, all projects must be supervised at the same level of importance.

- **Timeliness**

It refers to whether the data accurately reflects the current state of the system. It can be measured by the delay between a change in the real system and the associated change in the database, or by the frequency of data refresh. A delay of less than one hour is considered acceptable for real-time performance indicators (duration and completion rate data) because the frequency of use of the dashboard will be more than one hour. For other indicators, a delay of less than one day is acceptable as it will not affect the daily planning process.

- **Accuracy**

It refers to the degree to which the data correctly describes reality. It can be measured by comparing collected values to actual values. Accurate data present at least 90% of correct values. This level of accuracy ensures that any variations or fluctuations in the data are within a small margin of error, making it possible to identify trends and patterns in the data.

- **Consistency**

It is a measure of how consistently the data is recorded across different systems. It can be measured by determining whether two data elements representing the same aspect have any differences (format or meaning). If two elements of the same data have different format or meaning, the data is considered inconsistent.

- **Accessibility**

It refers to the ease with which data can be accessed, retrieved, and used by authorized users. This dimension is linked to SMART measurability criterion. Indeed, indicators are measurable if the data needed to track them is accessible. By keeping only measurable indicators (described in sub-section 3.4), it is reasonable to assume that the data is accessible. Therefore, there is no need to assess accessibility further.

Since there is a certain dependency between the aforementioned dimensions, we establish a hierarchy between them. Completeness has the top priority, followed by accuracy, consistency, and timeliness. As an example, when the data is incomplete, it is challenging and even impossible to test it against accuracy or consistency as there is not enough data to make such assessments. Similarly, when the data is inaccurate, it is difficult to know whether it is consistent or not, as the data does not reflect real-world values.

4 PRELIMINARY RESULTS

4.1 Performance indicators

At this stage of our study, we obtained results for the tactical dashboard indicators. Out of the 18 tactical performance indicators proposed (see sub-section 3.4), only eight were selected for developing the dashboard for the LIFE factory. Table 1 provides a brief description of each one of the selected indicators.

Table 1. Tactical performance indicators

Indicator	Definition	Reference
Order to delivery lead time	Total time from the beginning (order) to the end of the process (delivery)	[Khan and Bilal, 2019] [Tokola et al., 2016] [Kambartsumjan, 2021]
Production lead time	Time from the beginning to the end of production	[Tokola et al, 2016] [Kambartsumjan, 2021]
Delivery reliability	Ratio of orders delivered on time to the total orders	[Kambartsumjan, 2021] [Tokola et al, 2016] [Sipila, 2019]
Resource utilization rate	Actual manufacturing time vs. actual busy time of the machines	[Khan and Bilal, 2019] [Tokola et al, 2016] [Sipila, 2019] [Kambartsumjan, 2021]
Factory productivity	Total number of manufactured parts in a defined time	[Tokola et al., 2016] [Kambartsumjan, 2021]
Production rate	Ratio of the number of items manufactured to the average value	[Tokola et al., 2016] [Kambartsumjan, 2021]
Production process ratio	Ratio of actual production time to the order to delivery time	[Khan and Bilal, 2019] [Sipila, 2019]
Effectiveness	Ratio of actual production time to the estimated time	[Khan and Bilal, 2019] [Sipila, 2019]

According to [Siong and Eng, 2018], the most common ETO performance measure is delivery reliability, also called on-time delivery. It is related to lead time, which allows the managers to have more accurate predictions for delivery dates. Here we decided to track the order to delivery lead time, but also the production lead time because in the factory, the production time prediction is very important for planning.

The indicators related to productivity are used to compare the performance of the factory to previous results (factory productivity and production rate indicators) and to the factory's objectives (effectiveness). By tracking these indicators, the managers can identify areas of improvement and implement changes to increase productivity.

4.2 Data quality

The data collected from LIFE factory consists of several Excel files containing information on projects, tasks, machines, and employees over four years. We focused on the data available in two Excel tables: "Task" and "Project". That includes dates, completion rates, status of the operation steps, actual or estimated durations, etc. Each task from the Task table is associated with a project from the Project table (via a column called "unique project key"). These two tables were selected as they contain data required for the selected performance indicators. The remaining files contain mostly static data and were not included in the assessment. The assessment involved analyzing the data to identify any issues, and then, when possible, proposing and implementing measures to correct them.

4.2.1 Data quality assessment results

Table 2 presents the results of our data quality assessment. It shows only data categories with quality issues. A grey cell indicates that the data category does not meet the corresponding criterion. N/A stands for non-applicable, and NC for non-consistent. (T) and (P) mean that the data is from the Task table, and the Project table, respectively. Dimension names have been shortened to save space.

Table 2. Results of the data quality assessment

Data category	Compl.	Accur.	Consis.	Time.
Date_Debut (T)	100%	100%	NC	< 1h
Date_Fin (T)	100%	100%	NC	< 1h

Date_Debut (P)	26%	N/A	N/A	< 1h
Date_Fin (P)	23%	N/A	N/A	< 1h
Taux_Completion (P)	100%	87%	NC	< 1d
Date_Livraison (P)	87%	53%	N/A	< 1d

As an example, the data category "Date_Debut (P)" refers to a field in the Project table that provides the production starting date for a given project. Out of 647 projects (four years of data), only 170 have a starting date recorded (i.e., 26% of the projects). Therefore, Date_Debut (P) is not complete (since 26% is less than 100%). Regarding timeliness (for the same data category), the frequency of data refresh is less than a day as GestLIFE software automatically updates the date when a change occurs. It means that this data category meets the timeliness criterion. Unfortunately, we did not have enough data to assess accuracy and consistency. These criteria are non-applicable for this data category. The category "Date_Livraison (P)", which provides the project delivery dates, is also incomplete and inaccurate. Indeed, we identified from the data that 47% of the delivery dates recorded are before the end of production, which is impossible. It means that only 53% of the delivery date could be accurate, which is not enough to meet the accuracy criterion (90% of accuracy is required).
In addition, we identified that the completion rate of projects ("Taux_Completion (P)") was inconsistent. Further investigations revealed that the current method of calculating completion rates does not consider the assembly stage of production or the process of products order from external companies. The BPMN diagram already highlighted the lack of information storage regarding the assembly stage. The completion rate of a task makes sense, but, at the project level, which includes the process of ordering items from external companies or assembly stages, the current method used is inconsistent.

4.2.2 Corrective actions and discussion

To address the data quality issues identified through our analysis, we have proposed and implemented several corrective actions, mostly calculation-based, as the aim of the project was to use only available data. First, we calculated the production start and end dates of the projects for which the information was incomplete. To this end, we used the dates of the tasks associated with the projects. We took the earliest starting date of all related tasks as the production start date of the project, and the latest end date as the production end date. This allowed us to have dates for all projects. We also estimated the completion rate of the projects based on the durations associated with the related tasks.
However, the problem associated with the assembly stage (and external orders) is still not solved. We decided to remove the indicators affected by this issue until the factory could collect accurate data. Without this data, it is not possible to measure the order delivery lead time or the delivery reliability. Therefore, out of the eight tactical performance indicators presented in Table 1, we can currently measure:

- Production lead time
- Resource utilization rate
- Factory productivity
- Production rate
- Effectiveness

We recommend that the factory implements corrective actions to improve its management system (GestLIFE) such as considering the assembly process, external orders, and the remanufacturing of defective items. This would enable the collection of data on the entire process and help improve delivery-related performance. In addition, we suggest maintaining contact with the client after product delivery to measure, for example, customer satisfaction as part of the product quality performance.
A project is currently underway to implement a new production management software at the LIFE factory that would generate more accurate data on the manufacturing process. Beside implementing the recommendations outlined above, the LIFE factory can take considerable steps towards achieving its objectives of reducing lead time and maintaining product quality.

5 CONCLUSION

This study aims to develop a dashboard for the ÉTS Institutional Manufacturing Laboratory (LIFE) in order to support its transition into a smart factory. We found that only a limited number of indicators could be measured. This was due to a combination, among others, of unresolved data quality issues and a lack of data. We identified that order delivery lead time cannot be measured without accurate delivery dates while the quality of the products cannot be measured without additional data. In conclusion, the digital transformation of a factory and the implementation of a dashboard to monitor its performance engenders a need for high-quality data, and therefore, a rigorous analysis of data quality is essential to ensure the accuracy and reliability of the performance indicators used in the dashboard.
In our study, we used only the data already available within the factory. However, it is important to note that data collection can be further optimized by identifying specific areas where additional data collection is necessary and further data acquisition methods are required. The use of new technologies for data collection can greatly support this process. With the advancements in IoT and other Industry 4.0 technologies, it is possible to collect high-quality real-time data from various sources such as machines, products, etc.
In our work, we focused mostly on interviews with managers (and surveys) to select performance indicators, but we did not use a specific framework such as Goal-Directed Task Analysis to elicit information requirements [Nasser-dime et al., 2021]. This could be investigated in future work. Finally, due to time limitation, we could not collect the feedback of all the employees, and we did not survey the clients, which would be useful to measure client satisfaction. It can be another interesting improvement avenue in the future.

6 ACKNOWLEDGEMENTS

We would like to express our gratitude to the employees and managers of LIFE factory, particularly Mr. Xavier Rimpault and Mr. Eric Marcoux, for their contribution to this study. This research project is funded by the MEIE (Ministère de l'Économie, de l'Innovation et de l'Énergie du Québec) - OTN (Offensive Transformation Numérique) Program.

7 REFERENCES

Rochas, A., Boukherroub, T., & El Asli, N. (2022). Les expériences de gestion de la transformation numérique dans le domaine de la santé: Une revue de la littérature. *11ème conférence francophone en gestion et ingénierie des systèmes hospitaliers GISEH'22. 6-8 Juillet, Saint-Étienne, France*.

Few, S. (2007). Dashboard confusion revisited. *Visual Business Intelligence Newsletter*.

Eckerson, W. W. (2011) *Performance Dashboards: Measuring, Monitoring, and Managing Your Business,* 2e édition, John Wiley & Sons.

Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What Do We Talk About When We Talk About Dashboards? *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 682-692.

Rahman, A. A., Adamu, Y. B. & Harun, P. (2017). Review on dashboard application from managerial perspective. *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*.

Larsson, D., Ratnayake, R. M. C., & Gildseth, A. (2021). Development of a Balanced Score Card for Knowledge Work in Project-oriented Engineering Organization: KPI Prioritization Using AHP. *2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*.

Tokola, H., Gröger, C., Järvenpää, E., & Niemi, E. (2016). Designing Manufacturing Dashboards on the Basis of a Key Performance Indicator Survey. *Procedia CIRP*, *57*, 619–624.

Cerquitelli, T., Pagliari, D. J., Calimera, A., Bottaccioli, L., Patti, E., Acquaviva, A., & Poncino, M. (2021). Manufacturing as a Data-Driven Practice: Methodologies, Technologies, and Tools. *Proceedings of the IEEE*, *109*(4), 399–422.

Williams, D. & Tang, H. (2020). Data quality management for industry 4.0: a survey. *Software Quality Professional*, *22*.

Cichy, C. & Rass, S. (2019). An Overview of Data Quality Frameworks. *IEEE Access*, *7*, 24634-24648.

Ly, R. (2020). *Data quality management in data integration* [Master thesis]. Masaryk University.

Object Management Group (2011). *Business Process Model and Notation (BPMN), Version 2.0, formal/2011-01-03.*

Hicks, C., McGovern, T. H., & Earl, C. (2000). Supply chain management: A strategic issue in engineer to order manufacturing. *International Journal of Production Economics, 65(2),* 179–190.

Khan, M. A. R., & Bilal, A. (2019). Literature Survey about Elements of Manufacturing Shop Floor Operation Key Performance Indicators. *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*.

Kambartsumjan, A. (2019). *Designing, implementing, and evaluating a real-time performance dashboard* EMaster Thesis]. University of Twente.

Sipilä, A. (2019). *Analysing production flow of discrete manufacturing systems using simple node-based data* [Master thesis]. Tampere University.

Siong, B. C., & Eng, C. K. (2018). Implementing Quick Response Manufacturing to Improve Delivery Performance in an ETO Company. *International Journal of Engineering & Technology*, *7(2.28)*, 38.

A. Nasser-dine, A. Moïse and J. Lapalme, "Does Explicit Categorization Taxonomy Facilitate Performing Goal-Directed Task Analysis?," *in IEEE Transactions on Human-Machine Systems, vol. 51, no. 3, pp. 177-187, June 2021.*