Journal of Manufacturing Systems 58 (2021) 168–179

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

A knowledge-based Digital Shadow for machining industry in a Digital Twin perspective

Asma Ladj [a], Zhiqiang Wang [a], Oussama Meski [a], Farouk Belkadi [b], Mathieu Ritou [a,*], Catherine Da Cunha [b]

[a] LS2N (Laboratory of Digital Sciences of Nantes, UMR CNRS 6004), University of Nantes, Nantes, France
[b] LS2N (Laboratory of Digital Sciences of Nantes, UMR CNRS 6004), Centrale Nantes, Nantes, France

ARTICLE INFO

Keywords:
Digital shadow
Digital twin
Data and knowledge management
Machining

ABSTRACT

This paper addresses the problems of data management and analytics for decision-aid by proposing a new vision of Digital Shadow (DS) which would be considered as the core component of a future Digital Twin. Knowledge generated by experts and artificial intelligence, is transformed into formal business rules and integrated into the DS to enable the characterization of the real behavior of the physical system throughout its operation stage. This behavior model is continuously enriched by direct or derived learning, in order to improve the digital twin. The proposed DS relies on data analytics (based on unsupervised learning) and on a knowledge inference engine. It enables the incidents to be detected and it is also able to decipher its operational context. An example of this application in the aeronautic machining industry is provided to stress both the feasibility of the proposition and its potential impact on shop floor performance.

1. Introduction

The proliferation of Information and Communication Technologies (ICT) has led to the emergence of intelligent manufacturing paradigms. These emphasize the utilization of various methods, i.e. digital modeling, simulation, and experimental verification to control product design, resource allocation and production process at workshop level [1]. The new era that combines virtual reality technology based on the Cyber-Physical System (CPS) and the internet of things is then reached [2,3]. Digital Twin technology is the core component of CPS that is able to sense and reflect accurately the behavior and real-time state of the production system [4]; so that processes can be analyzed, simulated, predicted and optimized [5].

These technologies enable a large volume of data from various manufacturing activities to be collected and managed. In most cases, these data are spread out in different storage locations and are rarely shared between departments. So, the data collected in factories nowadays remain underexploited and generally used only by a few actors.

Faced with this challenge, the Digital Twin relies on a core component, called the Digital Shadow that enables the management and analysis of near real-time data coming from the actual physical counterpart [6]. It also relies on various simulation and visualization components. Theoretically, the Digital Shadow aims to achieve a comprehensive structuring of heterogeneous kinds of data available on the manufacturing shop floor. The data is then connected to their respective semantics in order to facilitate retrieval, interpretation, and exploitation. The role of the Digital Shadow of a production system is to enhance the digital image of the machines and assembly stations in the factory which uses information technologies [7].

However, several barriers are stopping the development of the Digital Twin and Shadow concepts in a real industrial environment. Firstly, the variety of business processes and the heterogeneity of data types lead to interoperability issues and lack of communication between all decision and operation centers. This results in problems for gathering the relevant information, at the right moment, for a given decision in relation to the variety of experts' interests. Secondly, it is generally assumed that the digital and simulation models are pre-defined and used in a static way in the Digital Twin. Nonetheless these models are difficult to update regularly in order to represent the dynamic and changing behavior of the machine throughout its life cycle. The operator and engineer have already developed at machine level their best practices and knowledge in order to cope with the real machine behavior.

Consequently, the implementation of advanced technologies in the factory requires those in industry to be extremely involved. It is

* Corresponding author.
E-mail address: mathieu.ritou@univ-nantes.fr (M. Ritou).

https://doi.org/10.1016/j.jmsy.2020.07.018
Received 30 October 2019; Received in revised form 22 July 2020; Accepted 24 July 2020
Available online 5 August 2020

important, therefore, to use an efficient knowledge engineering approach throughout the digital transformation process so as to integrate the experts' intention and viewpoint in data processing and interpretation. The originality of the proposed framework is to combine traditional data-driven approaches, generally based on data mining and statistical methods, with a knowledge-driven approach. The business experts' rules are extracted and inferred to interpret and augment the results of data processing, enabling the computation of new key performance indicators for a variety of business perspectives.

To do so, this paper proposes a knowledge-based Digital Shadow, as a core component and the backbone of future Digital Twin applications in manufacturing. Based on structured and modeled knowledge, an unsupervised machine-learning method is carried out in order to process the collected data and decipher the operational context. This is done so as to generate new knowledge regarding machining incidents. A knowledge-engineering method is used to capitalize (in the form of business rules) and reuse useful insights learned from past experiences and/or data analysis results. The problems of structuring, processing and exploiting the large volume of manifold data in production are solved and the integrated data-knowledge loop is closed.

This paper is structured as follows. In Section 2, applications of Digital Twin and Digital Shadow in the context of Smart Manufacturing are reviewed to clarify the positioning of each concept and to highlight the use of Digital Shadow for Digital Twin applications. In Section 3, the proposed framework of Digital Shadow is introduced. Firstly, data and knowledge models are explained. Secondly, the method developed for data analytics is described. In Section 4, a case study of machining aeronautic parts is carried out to demonstrate the feasibility of the approach. Finally, in Section 5, a brief conclusion and ideas for further development are presented.

2. Literature review

The in-depth integration between these smart devices in the physical world and their cyber environment is widely recognized as a key feature of smart manufacturing [8,4]. Relating to the production context, Cyber-Physical Production Systems (CPPS) consist of "autonomous and cooperative elements and subsystems that are connected within and across all levels of production, from processes through machines up to production and logistics networks" [9].

The CPPS technologies wisely combine data acquisition and transmission in physical space with data analysis in cyberspace through the 3C technologies (computation, communication, and control) [4]. Thereby, the physical entity can provide data to update its virtual model to achieve high flexibility, while a digital mirror model in cyberspace can be used to monitor, control and send global feedback to the physical entity to achieve high efficiency [52]. This smart process is managed at the top level by the concepts of digital twin and digital shadow.

The Digital Twin is applied in smart cities, healthcare, agriculture, the automobile industry, aerospace, manufacturing, etc. [10]. In the following sections, the application of Digital Twin as well as Digital Shadow in manufacturing will be discussed.

2.1. Digital twin for smart manufacturing

The commonly recognized definition of Digital Twin was given by Glaessgen and Stargel [11] who defined Digital Twin as *"an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, and so forth, to mirror the life of its flying twin."*

It is worth noting that Digital Twins are more than just pure data, they also include models and algorithms, which ensure a maximum concordance between physical and virtual spaces, as well as retroactive actions (feedback) to the physical system. In fact, beyond system simulation at the early planning stage, a Digital Twin enables a relevant use of the simulation during the system run-time [12]. Moreover, Digital Twin can be used for monitoring, control, diagnostics, and prognostics [13]. A Digital Twin reference model was proposed by Lu et al. [5]. It consists of three elements: (i) real (physical) space, (ii) virtual (digital) space and (iii) the two-way communication to reflect the dynamic mapping between them, i.e., data flow from real space to virtual space and information flow from virtual space to real space.

From a manufacturing perspective, ISO/DIS 23247-1 defines Digital Twin as a living model of manufacturing elements such as personnel, products, assets, and process definitions that both updates and changes as the physical counterpart changes [14]. Thereby, three main Digital Twin types are highlighted: Digital Twins for *products*, Digital Twins for *manufacturing assets* and Digital Twins for *manufacturing processes* [15]. In addition, Lu et al. [5] mentioned Digital Twin for *people*. Digital Twin research has mainly focused on manufacturing assets: 85 % of prior Digital Twin applications are developed for manufacturing assets as outlined by a recent review on Digital Twin applications in smart manufacturing [5].

For instance, Digital Twins for *products* are constructed to simulate and monitor the behavior and state of the physical product during the production phase, as well as during the life cycle [8,16]. Digital Twin for *manufacturing assets* corresponds to the digitalization of physical execution units for a workshop, such as machine tools, cutting tools, and other resources. The properties of the manufacturing system (i.e. geometrical structure, material properties, process parameters, working status, operating and environmental conditions, etc.) are expressed using an integrated multi-dimensional model [15]. Through physical asset digitalization, manufacturers can gain a clearer picture of real-world performance and operating conditions of a manufacturing asset, such as comprehensive and real-time monitoring of machine-tool status and visualizing the machining process [17–20], fault diagnosis of rotating machinery [21,22], life prediction of machine-tool [23], smart process planning [24], scheduling optimization [25], etc. Moreover, Digital Twin for assets can provide a tutoring service, augmented assistance [22] and decision-making support for humans [17].

A *process* Digital Twin is required to connect the product design and manufacturing processes [15,26]. The Digital Twin of a *manufacturing process*, also referred to as Digital Twin for *factories* [5], must function as one single, complex system composed of several synchronized sub-systems [15]. It might include all Digital Twins involved in the respective process, i.e., *product* Digital Twin, *physical resources* Digital Twin, and *human resources* Digital Twin of the considered production line [27]. Real-time controlling instructions are fed back to the physical space to optimize the management of the process [27], support decision-making in the system design and solution evaluation [28], etc.

The Digital Twin relies on the Digital Shadow, which digitizes real processes to create a duplication of reality as identical as possible using process modeling and simulation [29].

2.2. Digital shadow for smart manufacturing

According to the level of data integration between the physical and digital counterpart in Digital Twin applications, three subcategories are considered: Digital Model, Digital Shadow, and Digital Twin [30]. *Digital Model* refers to the digital representation ensuring the comprehensive description of the physical counterpart, e.g. simulation models, mathematical models, etc. Digital Models do not use any form of automatic data exchange with their physical objects. If an automated one-way data flow directed from the physical object to its digital object exists, it is a *Digital Shadow.* A change in the physical object state leads to a change of the digital object. Furthermore, the data flows are fully integrated for both directions, one may refer to *Digital Twin* where the digital objects generate feedback for the control of the physical object [30]. Hence, a Digital Twin contains all knowledge resulting from modeling activities in engineering (Digital Model) and from working data collected during real-world operations (Digital Shadow). With appropriate simulation algorithms, it is possible to obtain an "Experimentable Digital Twin"

[31] as shown in Fig. 1. Thus, the physical system performance can be tested and evaluated under different boundary conditions and in various operational scenarios [32].

In the manufacturing context, every product or component produces a data profile (including operation and condition data, process data, etc.) that depicts a Digital Shadow. It is defined as an integrated database that generates a sufficiently precise digital representation of the production system in real-time [29]. The Digital Shadow is linked in real-time with the manufacturing system and generates a database for the optimization [4]). This database is created using algorithms that include the acquisition, analysis, evaluation, and consolidation of data. Data gathered from multiple sources are aggregated and stored in a shared platform built by the Digital Shadow enabling data to be linked to their appropriate context [33]. The Digital Shadow has one core function, namely the supply of the right information at the right time and in the right place [29].

Without a Digital Shadow, a missing link will appear between the digital factory and a holistic digital image of the production [7]. Hence, the Digital Shadow is recognized to be “a prerequisite for the application of methods and models of data analysis and evaluation in a manufacturing environment” [6]. A Digital Shadow also provides semantically adequate and context-aware data from production, development, and usage in real-time with an adequate level of granularity [34]. It contributes to the comprehensive structuring of different kinds of data collected from different data sources that can be stored in a shared repository and can be used for numerous applications [33]. A well-designed Digital Shadow forms a promising key for a comprehensive analysis of manufacturing systems. Furthermore, the Digital Shadow serves as a basis for further applications that use the aggregated data when it is no longer necessary to define interfaces between each application and its data sources because each application can refer to the structure of the Digital Shadow [33].

Schuh et al. [33] proposed a Digital Shadow framework that enables data collection and integration of the entire maintenance, repair, and overhaul services, regardless of the source of the data. Schuh et al. [6] developed a data structure model for the Digital Shadow which enables the efficient use of the knowledge management system to support companies of small batch production on the way toward Industry 4.0. In order to enable overarching explorative analytics in product life cycle

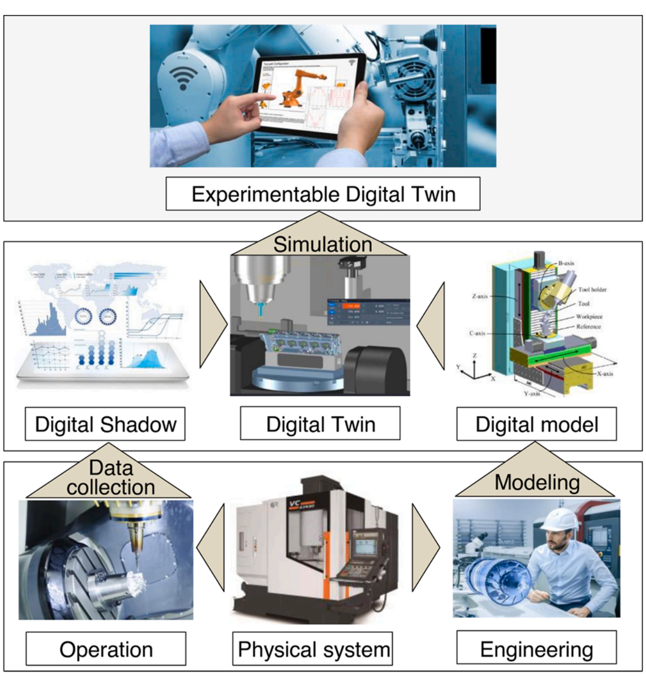


Fig. 1. Digital Twin, Digital Shadow, Digital Model concepts [31].

management, Riesener et al. [35] proposed a model that merges information based on heterogeneous data sources. Three steps were identified for acquiring the necessary information for a Digital Shadow and for choosing the best data source for generating this information. A data-information-fit-indicator was introduced to choose a suitable data source for the required information. An integration of Digital Shadow simulation model with the Manufacturing Execution System (MES) has been proposed by Negri et al. [36] with the aim of creating a Digital Twin. The decision making in the MES-integrated Digital Twin is optimized by using an intelligent layer that hosts the rules and the knowledge in order to choose from alternatives.

As a conclusion, new concepts and architectures were recently proposed for Digital Twin and Shadow, but their maturity is currently very low [5]. However, Digital Shadow plays a crucial role, particularly when there is complex data or big data.

3. Proposed framework of knowledge-based digital shadow

In this work, the focus is on the Digital Shadow: a unidirectional flow of data processing from the physical counterpart to the digital object is considered.

The physical space of the proposed framework consists of the physical system: the machine-tool and the smart sensing device. The properties of the machine-tool and process need to be obtained from diverse sources, ranging from the geometrical structure, work-piece material properties, process parameters, working status, to the operating status (cutting depth and speed, etc.) and environmental conditions (ambient temperature, etc.). Smart sensing techniques can be used to complete these attributes: accelerometers for vibration measurements, communication by field bus to the CNC, etc.

To design an efficient Digital Shadow for manufacturing, we have developed leveraging data and knowledge models. This integrates an information model that abstracts the specifications of the machine tool, as well as a shared knowledge base acquired either from the analysis of data in the form of thresholds or from the interaction with experts in the form of capitalized business rules. The knowledge base is continuously enriched by direct or derived learning. Then, a data analytics module is proposed to merge manifold data, check its consistency, correct errors, and process it using a machine learning means in order to extract useful information about the machine tool behavior and machining state. Finally, based on performed data, aggregated data can be generated to monitor the machine and machining states. This visualization helps in making initial decisions. The scheme of the proposed framework is shown in Fig. 2, where highlighted modules correspond to the present contribution; the remaining components are beyond the scope of this paper.

3.1. Data and knowledge models

The functioning of a Digital Twin and shadow is primarily based on the simulation and the processing of real data collected from the physical system. This data can vary depending on the nature of the system to be monitored, and the analysis/decision making objectives of the Digital Twin. The different data processing and aggregation algorithms explained above generate and use several types of data that should be coupled and organized correctly to make the computing process automatic. Consequently, database structure and content could cover product-oriented and/or process-oriented perspectives. Additional knowledge has to be stored and processed in order to support the interpretation and exploitation of the above data within their contexts of creation and use. For instance, learning algorithms will result in different rules that clarify the detection of undesirable events or a particular behavior.

Researchers are often interested in studying one or two categories of these data sources according to their development objectives. The work is, therefore, more complicated when considering a heterogeneous

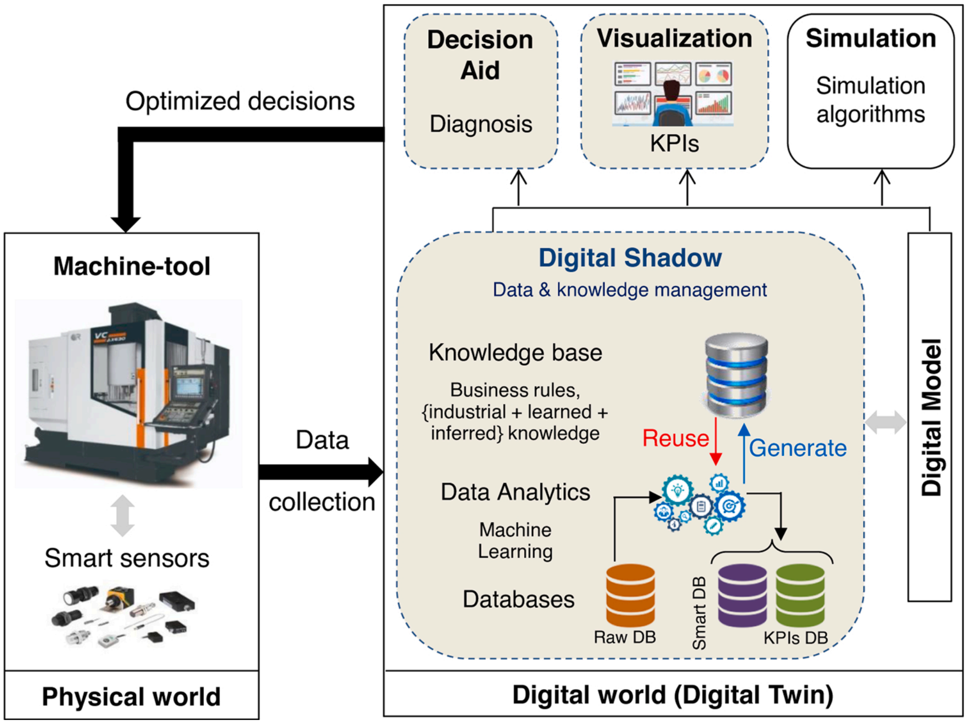


Fig. 2. Digital Twin global scheme highlighting the proposed framework.

collection of knowledge that comes from very different sources with very different frequencies and formats. Consequently, it is essential to use a structuring and modeling method for data and knowledge in order to create links between them and facilitate their manipulation. An important step to distinguish all the data and knowledge to be managed in the Digital Shadow/Twin repository consists of the analysis of the AS-IS situation of the industrial system and all related environments.

In the literature, research works that address the problems of Digital Twin and Digital Shadow implementation generally consider several types of models: data models, information models, business models, and service models that are supposed to cover the different categories of data (resources, processes, and outcomes) [37]. However, the existing models do not consider the different phases of the service. In particular, data related to performance monitoring and improvement, and those related to failure recognition, are often missing or lack details. Consequently, these models can endanger the implementation of a reliable Digital Shadow [38]. To overcome this limitation, a global knowledge management solution is developed in order to handle the entire heterogeneous knowledge available in the service. Thus, in the proposed framework, data and knowledge are structured using the inference model to develop an extensible Knowledge Base (KB), as shown in Fig. 3. The development consists of three stages; the first phase of industrial knowledge capitalization, which results in structuring the global knowledge base; for the second phase of reasoning that generates inferred knowledge and will be reintegrated into an ontology in a third phase to provide semantic-rich and interoperable models that allow reasoning and automatic inference, ontology-based models are used.

Noy and McGuiness [39] defined ontology as "A formal and explicit description of concepts in a discourse domain (classes or concepts), properties describing the attributes and characteristics of each concept (roles or properties) and restrictions". In the engineering field, ontology must represent not only product-related information, but also meta-information, which is a corollary to the product, the process, the organization, or even the resources and rules. The ontology is developed using "Protégé" software. It is open-source and free. It offers the possibility of integrating business rules to infer and also the possibility of communication with other reasoning software. Another advantage is the

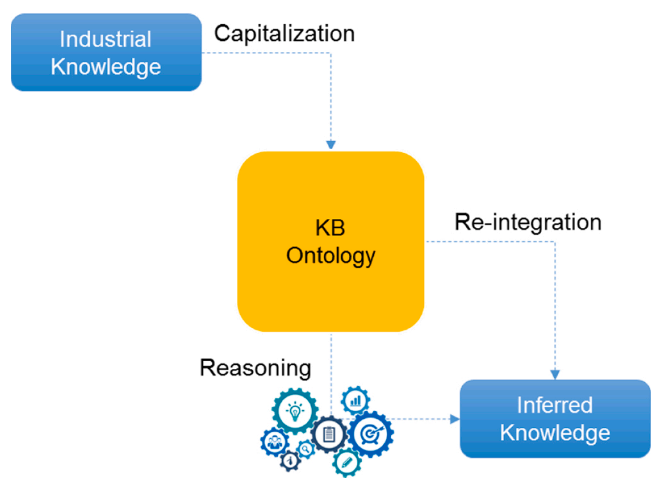


Fig. 3. The generic knowledge management approach.

possibility of integrating developed ontologies into frameworks, through the availability of several APIs that enable communication with different development tools and programming languages. This is very useful for research projects.

Fig. 4 was extracted from the development of the proposed ontology using “Protégé”. It shows the different elements required to create the ontology. Let us take the example of two *classes* "Product" and "Process". It is possible to define relationships between them by defining an *Object property* "HasManufProcess" and associating attributes for each class using the *Data property* "ProcessId" and "ProductSerialNum". Then it can be instantiated by adding *individuals*.

Ontologies can be classified into two main types: storage ontologies and inference ontologies [40]. The specificity of the first category is their ability to capitalize a maximum of knowledge from a specific domain or sometimes from several domains (in the case of hybrid ontologies, for example). The interest of the inference ontologies is to be able to deduce instance affiliations to classes thanks to the axioms and

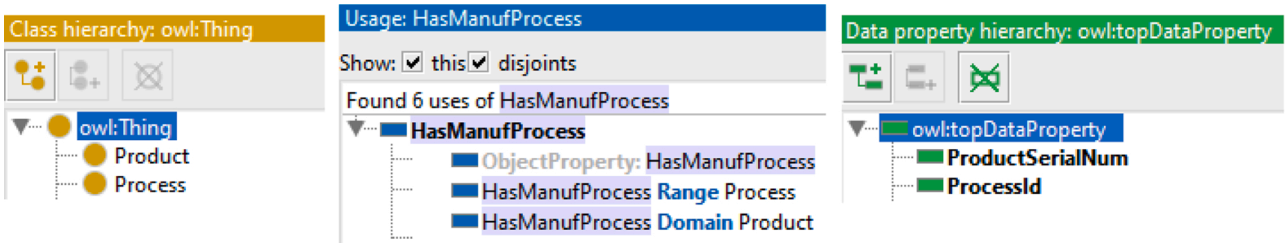


Fig. 4. Basic elements of an ontology.

restrictions of ontology [41].

Once the development of ontology has been achieved, a consistency check phase is required. For this stage, it is necessary to launch an instantiation stage which involves a data and knowledge capitalization process.

3.1.1. Data and knowledge capitalization

To develop a knowledge base, the first phase of capitalizing integrates a maximum of data and business knowledge. Two categories were considered: (1) norms and standards from the literature, as well as dictionaries about the different processes, products and resources used: namely machine-tools, cutting tools and their cutting conditions, etc. This first category enables the capitalization of the knowledge of the manufacturing domain. Norms and standards provide a generic aspect to the ontology which promotes its adaptability in other industrial applications and avoids differences of semantics between companies. (2) Context-dependent data and knowledge which represent the specific knowledge of the use case and which group together the knowledge of the industrial companies.

The second source relies on the analyses of the processes and industrial data of the user. This stage can be automated thanks to the implementation of a machining monitoring system. It follows several parameters and characteristics of the machine at the time of machining by using sensors that measure power, acceleration, temperature, etc. Given the variety of available sensors, the selection of the best according to application requirements and sensor specifications must be made wisely [42].

This monitoring operation generates a large volume of raw databases. Then, using data mining methods, machining data are analyzed and useful information is extracted [43,44]. Contextual classification and aggregation phases generate more meaningful data, called Smart Data, that can detect events that occur during machining [45]. Smart Data are then used to generate Key Performance Indicators (KPIs) that describe the machining process and its potential detrimental phenomena.

This research work provides a link between knowledge management and data analysis domains (which will be detailed in the following section). The use of ontology is a key point that establishes this link. The ontology represents the support necessary for the deployment of the different decision-aid algorithms. The first axis of decision-aid is reporting, which enables the generation and feedback of reports combining several KPIs that describe the malicious incident being detected. All elements required for this reporting process are available in the ontology. Among these elements, the ontology integrates concepts for the detection of a detrimental incident, such as tool failure, collision or chatter.

3.1.2. Business rules

Business rules can be formalized from either explicit or implicit knowledge. Explicit knowledge can be capitalized directly through the analysis of literature, for example technical documentation and user manuals. Moreover, implicit knowledge is often linked to the specific know-how of experts [46]. These rules allow, on the one, hand building and defining constraints on a specific process and, on the other hand, they enable a decision-making and process control in the company, based on a set of well-defined criteria. The development of business rules is divided into three stages: during the first stage, as much of the potentially recoverable expertise as possible is questioned. The information is then qualified and processed to check consistency and remove redundancies in the second stage. The third phase translates these rules into a language that can be understood and used by inference engines.

The global knowledge base also supports the reintegration of knowledge, given the nature of the developed ontology of "inference". The inferred knowledge is essentially generated using the business rules defined and validated in collaboration with the industrial experts. In addition to detection, the knowledge base can also be used for diagnostic purposes.

3.2. Data analytics

In the context of Industry 4.0, large volumes of manufacturing digital data are available on instrumented machine-tools. The challenge is to exploit them by mining to improve decisions.

Fayyad et al. [54] formalized the general approach of Knowledge Discovery in Database (KDD), in which data mining is one of the steps to discover interesting hidden information (pattern) in data. KDD is generic and applicable to any kind of data, regardless of the context or of previous knowledge. In an industrial context, if the business expertise is not integrated, complex investigations could lead to discovering only trivial, explicit or previously known patterns. The CRISP-DM [53] process is also generic and was developed to limit this risk and to support the implementation of the KDD approach in a company. It takes advantage not only of databases but also of the pre-existing business knowledge. Here, the six steps of CRISP-DM (Fig. 5) are followed. In addition, we propose specific monitoring criteria, business rules and KPI, managed by ontology.

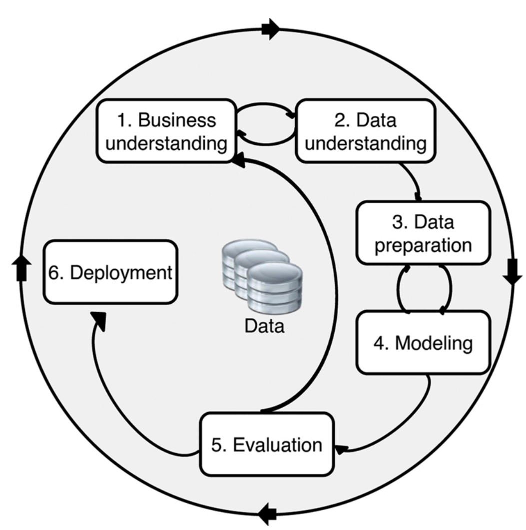


Fig. 5. The CRISP-DM life cycle [51].

3.2.1. Business understanding

The first step aims to determine the business objectives, identify the needs and the available resources. The proposed Digital Shadow includes business knowledge in this resources list, taking advantage of the company's expertise.

In our case, the business objective is the reduction of detrimental phenomena. The available business knowledge includes contextual information that enables us to:

- Describe the machine-tool working conditions (e.g., the day, program name, tool number, machine-tool motion, spindle motion, material removal).
- Collect the operational context when incidents (chatter, tool failure or collision) occur.
- Realize a more relevant calculation of Key Performances Indicators (KPIs).

In the contextual classification, it is interesting to know:

- If the spindle is stopped (S) or rotating at constant speed (CS) or varying speed (VS);
- If the machine-tool is stopped (S) or moving at constant (CS) or varying speed (VS);
- If the tool is cutting materials or not (N), with a constant (CC) or varying tool engagement (CV).

Seventeen potential states for the machine-tool can be proposed based on the combination of this contextual information. Table 1 presents these 17 states with their conditions. The objective is to determine the state of the machine-tool at each instant.

For example, in order to know if the tool is cutting materials or not, a classification according to the criteria Arms (root mean square-RMS of vibration acceleration value, in m/s^2) and the spindle power P (kW) is proposed, when the machine is moving (feedrate $Vf > 0$). In machining, it is known that if the tool is cutting, its vibration (Arms) is greater than if the tool is not cutting. It is also known that the power (P) will exceed the spindle power corresponding to idle rotation (P_{idle}) when the tool is not cutting. The proposed business rules are:

- ***If*** *($P > P_{idle}$or Arms $> T_{Arms}$) and Vf>0,* ***then*** *the tool is cutting material*;
- ***If*** *$P \leq P_{idle}$and Arms $\leq T_{Arms}$,* ***then*** *the tool is not cutting material.*

3.2.2. Data understanding

A sensing system collects the measured data during machining and stores it in a database. The data mining methods will then be used to recognize the noise data (e.g., abnormal values of accelerometers) and the missing data (e.g. loss of connection between CNC and the sensing system). The noise data will be eliminated or neutralized after the knowledge of its context. For example, the high value of accelerometers generated by lubrication could be neutralized because it is not the source of chatter. The missing data will be discarded.

3.2.3. Data preparation

Combing raw data and contextual information can help to detect, more precisely, the detrimental incidents (e.g., chatter during machining, tool wear, and collision). Finally, the data classified with the detected incidents are aggregated into simplified and more meaningful information by multi-level aggregation [45].

3.2.4. Modeling

This step aims to build a model, set its parameters, and assess its fit. In our case, unsupervised machine learning is performed on different descriptors (spindle speed N; feed rate Vf; spindle vibration Arms; spindle power P, etc.) to determine the classification thresholds (with which the different states of the machine-tool are determined). Each machine-tool has different characteristics, so it is necessary to perform initial training for each machine-tool. A data set corresponding to a period of one day of industrial production is used for this purpose.

The objective is to find, in the modeling step, the thresholds for cutting vibrations T_{Arms} (m/s^2) and for the idle spindle power P_{idle} that classify in the clusters 'tool cutting materials' or 'tool not cutting materials'. The learned values of thresholds for each descriptor are specific to a given machine-tool and should be capitalized in the knowledge base.

The learned thresholds are verified on a small time-interval: manual data mining is applied to check point by point. In the absence of classification errors on tool cutting or not, thresholds are validated.

3.2.5. Evaluation and deployment

The two last steps of the CRISP-DM method concentrate on supporting documentation of projects, capture experience for reuse, support knowledge transfer and training. In our case, once the model is validated, automated prevention of detrimental phenomena can be implemented into the machine-tool. The criteria (Nh, Ub, Vrms, A_{peak}) corresponding to different events are computed at each instant; if an incident occurs, the corresponding criterion should be higher than for normal conditions.

The actual impact of this modification has to be monitored and if proven successful the same approach can be applied to other machine-tools. Furthermore, the acquired assets have to be integrated into the knowledge base and shared between the different services (design, maintenance, etc.)

Table 1
Classification of the different states of the machine-tool.

<table>
<tr><th colspan="3" rowspan="5"></th><th colspan="7">Feedrate of the machine-tool Vf</th></tr>
<tr><th rowspan="4">$Vf = 0$</th><th colspan="6">$Vf \gt 0$</th></tr>
<tr><th colspan="3">$\Delta Vf \gt T_{\Delta Vf}$</th><th colspan="3">$\Delta Vf \leq T_{\Delta Vf}$</th></tr>
<tr><th rowspan="2">$Arms \leq T_{Arms}$ & $P \leq P_{idle}$</th><th colspan="2">$Arms \gt T_{Arms}$ or $P \gt P_{idle}$</th><th rowspan="2">$Arms \leq T_{Arms}$ & $P \leq P_{idle}$</th><th colspan="2">$Arms \gt T_{Arms}$ or $P \gt P_{idle}$</th></tr>
<tr><th>$\Delta P \leq T_{\Delta P}$</th><th>$\Delta P \gt T_{\Delta P}$</th><th>$\Delta P \leq T_{\Delta P}$</th><th>$\Delta P \gt T_{\Delta P}$</th></tr>
<tr><td rowspan="3">Spindle speed N</td><td colspan="2">$N = 0$</td><td>Spindle: S
Machine: S
Cutting: N</td><td colspan="3">Spindle: Stopped (S) Machine-tool: Varying Speed (VS) Cutting : No (N)</td><td colspan="3">Spindle : Stopped (S) Machine-tool : Constant Speed (CS) Cutting : No (N)</td></tr>
<tr><td rowspan="2">$N \gt 0$</td><td>$\Delta N \gt T_{\Delta N}$</td><td>Spindle : VS
Machine : S
Cutting : N</td><td>Spindle : VS
Machine : VS
Cutting : N</td><td>Spindle : VS
Machine : VS
Cutting : eng. Constant (CC)</td><td>Spindle : VS
Machine : VS
Cutting : eng. Variant (CV)</td><td>Spindle : VS
Machine : CS
Cutting : N</td><td>Spindle : VS
Machine : CS
Cutting : eng. Constant (CC)</td><td>Spindle : VS
Machine : CS
Cutting : eng. Variant (CV)</td></tr>
<tr><td>$\Delta N \leq T_{\Delta N}$</td><td>Spindle : CS
Machine : S
Cutting : N</td><td>Spindle : CS
Machine : VS
Cutting : N</td><td>Spindle : CS
Machine : VS
Cutting : CC</td><td>Spindle : CS
Machine : VS
Cutting : CV</td><td>Spindle : CS
Machine : CS
Cutting : N</td><td>Spindle : CS
Machine : CS
Cutting : CC</td><td>Spindle : CS
Machine : CS
Cutting : CV</td></tr>
</table>

4. Case study in the aeronautic machining industry

The proposed framework of Digital Shadow has been implemented in an industrial company that produces structural aeronautic parts in aluminum alloy. The specificity of such parts results from their high added value, requiring high-speed machining. This is characterized by high material removal rate, and thin walls and floors subject to vibration issues. Hence, the manufacturing process should be very accurately planned and produce good parts from the first one. In such a complex context, a Digital Shadow is necessary to efficiently exploit the in-process collected data and available knowledge for the monitoring of the machine-tool state, the machining process progress and quality.

4.1. Implementation

An in-process data collection system EmmaTools was installed on a machine-tool of an aircraft manufacturer factory as shown in Fig. 6. It collects data every tenth of a second. The data comes from two information sources: the Computer Numerical Control (CNC) and the added sensors. By field bus, the CNC provides machining context information such as the tool name, the current program, machine and spindle motions, etc. The sensors integrated into the machine-tool collect the instantaneous power of the spindle, the temperatures, as well as the vibrations from 4 accelerometers integrated into the spindle (radially at each bearing).

In order to detect incidents dedicated monitoring criteria are used. They are based on mechanical phenomena for better reliability. They are computed from real-time signals and collected in the PostgreSQL database. They are listed below:

- The spindle condition is evaluated through the monitoring of bearing fault-induced vibration BPFO (i.e. the amplitude of the vibration contribution at the Ball Pass Frequency of Outer ring, [47]), which is computed from the daily vibratory signature.
- Chatter (which is an unstable cutting phenomenon that results in unacceptable surfaces of the workpiece) is monitored by Nh (sum of the amplitudes of the five dominant non-harmonic contributions of the vibration spectrum), introduced by [43].
- Since tool failure leads to an increased mechanical unbalance of the cutting tool, it is monitored by the criterion Ub, proposed by [43], which evaluates the amplitude of the vibration at the spindle rotation frequency.
- The collision between any parts of the machine, work-piece, or clamping is very dangerous. The collision is monitored by the criterion A_{peak} (maximum value of the raw acceleration signal over a given period) [43].

4.2. Knowledge model

The specificity of this research work is the implementation of a Knowledge Management (KM) approach at the service of data analysis tools in order to facilitate the management of industrial data and knowledge flows. It is gathered in a generic knowledge repository, structured by ontology. The development of a knowledge base with a generic ontology facilitates the implementation of a decision aid system. The approach is presented in Fig. 7 and explained below.

The proposed KM approach covers the whole process of analysis and detection of malicious incidents that may appear during the machining process. Firstly, business experts are interviewed in order to formalize their expertise into business rules, regarding, for example, the detection of malicious incidents. Then, using an initial data base, detection thresholds are obtained by unsupervised Machine Learning (as explained in Section 4.3). Thresholds and business rules are capitalized in the knowledge base.

Thereafter, a reasoning system is set up to ensure communication between the data and knowledge bases. This system enables the detection of the occurrence of any problems during machining, once the dedicated business rule and detection threshold are retrieved from the knowledge base. The contextual information of the incident is also obtained by data mining and deciphered. In this way, beyond a simple time stamping of the incident, it is possible to know that the incident occurred, e.g. during a plunging operation of a given cutting tool, at a given feedrate in a specific workpiece program. This operational context and the detected incident are then instantiated in the ontology. It constitutes a first level of traceability. This approach facilitates communication between the sub-systems of the global framework and facilitates the management of available industrial data and knowledge, as well as decision support operations.

Concerning the use-case developed in Section 4.4, the objective is to detect a tool failure. The first step is thus to learn the threshold T_{Ub} and to store it in the knowledge base. Afterwards, and after recovering the values of the monitoring parameter *Ub* and the detection threshold, a reasoning based on business rules is triggered. In the example, the following rule was used:

If *(Ub* $>$ T_{Ub}*) & NotCutting* ***then*** *'tool failure detection'* $= 1$

The business rules are developed under the Semantic Web Rule Language (SWRL). Here, it is able to check the instances of the Ub parameter, the comparison with the detection threshold T_{Ub}, in relation to the machining state (*Cutting* or *NotCutting*), which is calculated using another business rule (defined in Section 3.2.1). Once detected by a monitoring criterion and its associated business rule, an incident and its operational context (tool reference, program, cutting conditions, machine state, etc.) is stored in the knowledge repository.

The use of a KM approach is very beneficial compared to traditional monitoring and detection systems and has several advantages. Firstly, the dynamic aspect of offline learning of the detection thresholds ensures the adaptability and genericity of the proposed solution. It can be transferred to other systems without necessarily needing to develop a specific system tailored for each use. Furthermore, the approach is highly useful in the case of handling heterogeneous knowledge. And

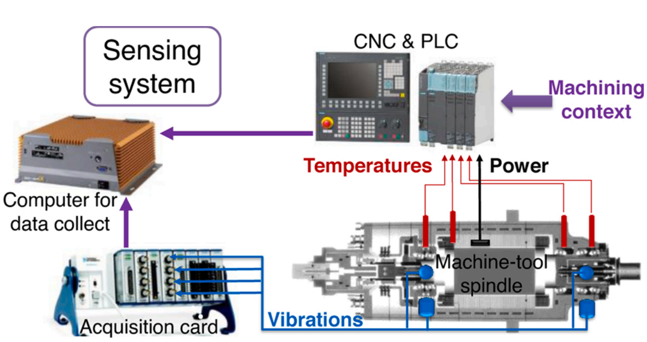


Fig. 6. In-process data collection system.

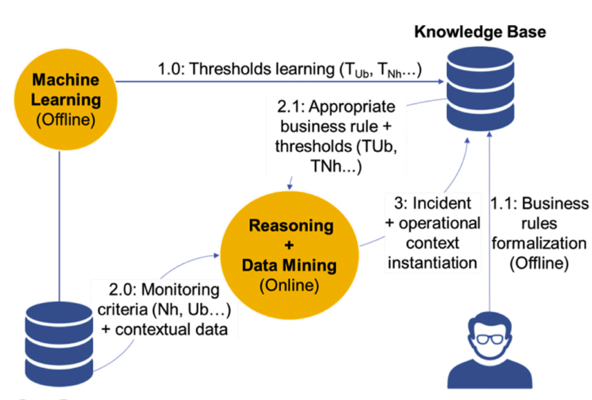


Fig. 7. The detailed knowledge management approach.

finally, a remarkable advantage of KM integration is the facilitation of the prospects of implementing decision-support scenarios.

4.3. Learning the business rule thresholds

The unsupervised machine learning by Gaussian Mixture Model (GMM) is used in this article. The Gaussian Mixture Model is a statistical model used to parametrically estimate the distribution of random variables by modeling them as a mixture of Gaussian distributions [48]. According to the contextual classification of business rules, the GMM is applied to the collected data in the context of 'spindle rotation at constant speed' (because in aeronautical manufacturing, the tool cuts materials only after the spindle rotates at constant speed) and of machine axes motion ($Vf > 0$). Three Gaussians can model the distribution of Arms in this class (N constant and $Vf > 0$) (interpreted as: 'the tool is not cutting' – Y1 green; 'the tool is cutting' – Y3 blue; and 'severe machining vibrations' – Y2 red). Finally, the distribution of Y4 (in cyan) is the sum of these 3 Gaussians (Y4=Y1+Y2+Y3). The threshold T_{Arms} is defined as the intersection between the two Gaussians Y1 and Y3, in order to minimize classification errors (false positives and negatives) (Fig. 8-a).

Three Gaussians can model the distribution of P in the class (N constant and $Vf > 0$) (that can be interpreted as '*the tool is not cutting*' – Y1 green; '*tool is cutting with low material removal*'– Y3 blue; and '*the tool is cutting with high material removal*' – Y2 red). The distribution of Y4 (cyan) is the sum of these 3 Gaussians (Y4=Y1+Y2+Y3). The threshold P_{idle} is identified at the intersection between the two Gaussians Y1 and Y2 (Fig. 8-b).

Thus, for data where '*spindle is rotating at constant speed*' and '*axes are moving*' ($Vf > 0$), if $Arms > T_{Arms}$ or $P > P_{idle}$, the data is assigned to the class '*tool is cutting*'. To check the classification by manual data mining, N (kRPM), Vf (m/min), Arms (m/s^2) and P (kW) are plotted in Fig. 9 (cyan for Vf, blue for N, pink for Arms). The results of the classifications concerning the tool engagement are plotted on the power P curve with different symbols and colors. The '*tool not cutting*' class is represented by black "+" signs on the power P curve; the '*tool cutting*' class is presented by red and blue dots. The '*tool cutting*' class can be classified into two sub-classes: '*tool cutting with constant engagement*' (in red points) and '*tool cutting with varying engagement*' (in blue points). When the spindle has reached its target speed (N = 23,720 rpm, blue lines), the tool is not cutting immediately.

Business knowledge enables to identify the existence of a subclass called '*spindle startup*'. This class is represented by green "+". It is defined by the following rule: from the reach of the target speed N until the power exceeds the no-load power P_{idle} (which represents the beginning of machining). The aim is to eliminate noise in vibration signals related to a lubrication issue during spindle startup which would otherwise lead to classification errors. The subclass '*spindle startup*' is also assigned to the cluster '*tool not cutting*'. This business rule facilitates the data analysis, by narrowing the clusters before data analysis.

The proposed unsupervised learning method by GMM applied to the classification on '*tool cutting*' *or* 'tool *not cutting*' was evaluated in [49, 50] where a confusion matrix revealed an accuracy of 99.94 %. The classifications of whether the tool is *cutting materials* or *not* are correct, which validates the proposed business rules and the threshold learning by GMM. The classification is based on the business rules and the GMM method. The validation of GMM can refer to Maugis et al. [48]. The learned values of thresholds for each descriptor are specific to a given machine-tool, and should be capitalized in the knowledge base.

In order to detect a tool failure, the criterion Ub is computed. It is the vibration amplitude of the contribution at the spindle frequency when the spindle is rotating without cutting (to prevent false detection due to cutting forces). Fig. 10 presents a histogram of Ub criteria for tool breakage, considering the whole spindle lifetime (215 days). The distribution of Ub was modeled using GMM with two components (normal population in green and tool failure population in red). The critical threshold T_{Ub} is then chosen as the abscise value where the two Gaussians cross (in order to minimize classification errors). A threshold T_{Ub} of 6.32 m/s^2 was learned, as shown in Fig. 10.

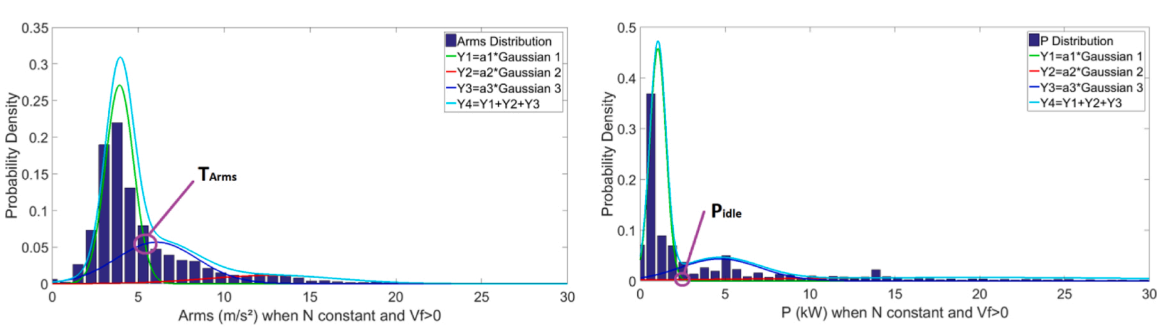
Therefore, the ontology classes "Ub" and "threshold T_{Ub}" are instantiated (see Fig. 13). Using a specific business rule and the object property "has_threshold", new knowledge about the "tool failure" is inferred.

4.4. Detection of tool failure

The tool failure detection is vital because the machine-tool could continue machining with a broken tool, damaging the work-piece and even the machine-tool. The criteria Ub is computed at each instant when the spindle is rotating without cutting. Data mining was performed in the database corresponding to one year of industrial production. Several detrimental incidents were found and one typical tool breakage will be illustrated.

By basic data mining, it can be found that large vibrations (Arms, in red) occurred during the 136th day of production, at the 6th hour in Fig. 11. However, without contextual information, machine learning and knowledge management, it is not possible to determine if it is normal or dangerous, nor to understand what really happened that day. This is the reason why a Digital Shadow is required, in order to enable efficient management and processing of the collected data and knowledge.

Fig. 12 presents the Ub criterion calculated between the 5th and 8th hours. The blue line plots Ub for tool breakage monitoring from the accelerometer signals. The dark line indicates the learned threshold T_{Ub} (from Fig. 10). As explained in Section 4.2, the business rules are processed by the knowledge inference engine, which compares the monitoring criterion Ub to the detection threshold T_{Ub} (this was obtained by unsupervised machine learning, cf. Fig. 10) in the context of *NotCutting* state (in order to avoid false alarms due to the cutting forces). The



(a) Distribution of Arms (b) Distribution of P

Fig. 8. Distribution of Arms and P (when N constant and $Vf > 0$) modeled by GMM.

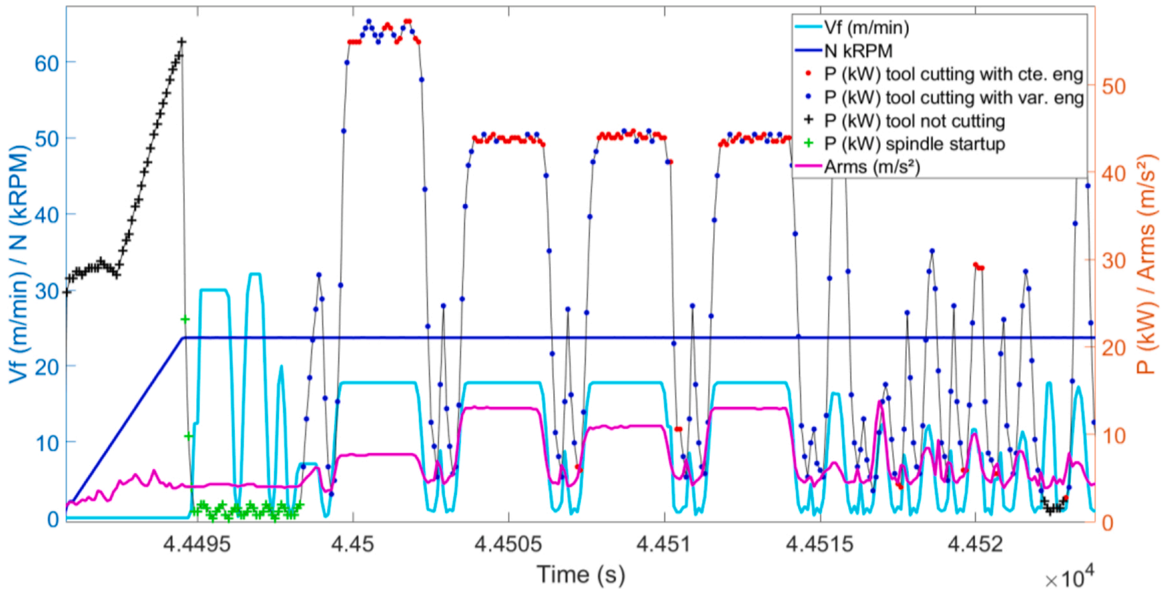


Fig. 9. Classifications of material removal by the cutting tool.

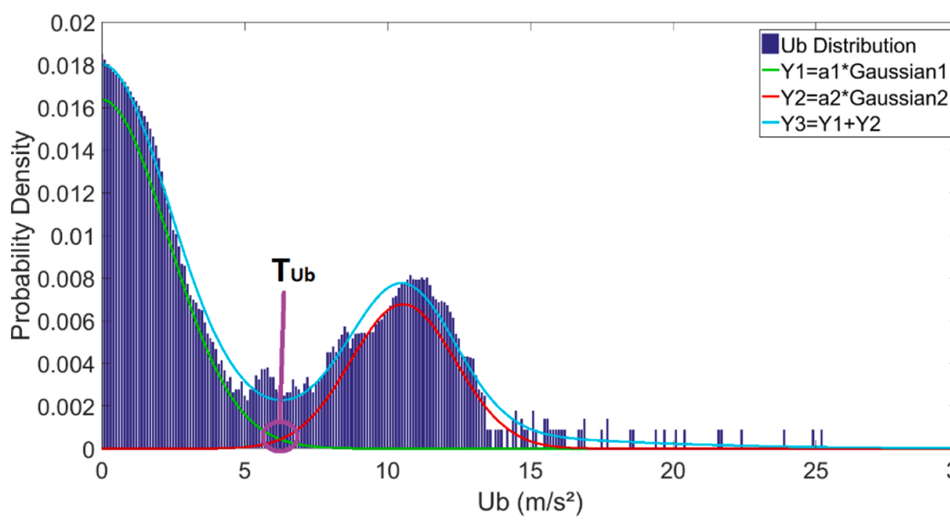


Fig. 10. Modeling of Ub distribution by GMM (composed of a normal population in green and of a tool failure population in red) and learning of the detection threshold T_{Ub} for tool breakage (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

execution of the rules generates a new direct knowledge: the instance of tool failure detection and its operational context which are stored in the knowledge repository. The instantiation of the incident is associated to the operational context, which means that additional elements managed by the ontology are instantiated: they are issued from the collected data (such as the tool reference 36 and its 5th instance, the program number 116, the day number 136, the cutting conditions, etc.) or from data analysis through the contextual classifications of the state of the process and machine-tool (cf. Fig. 9). In this way, the breakage of tool 36 is detected from the 31st to the 70th minute. The red line presents the Boolean for tool failure detection (1 – tool failure, 0 – tool in good condition). The green line shows the tool instance index. It can be observed that the machine-tool has changed the cutting tool (reference 36) at the 75th minute.

Contrary to the proposed Digital Shadow, the machine-tool had not detected this tool breakage before and has continued machining for 39 min with a broken tool, which is very dangerous. In this case, the Digital Shadow would have immediately detected the tool failure, enabling the machine-tool to stop thus avoiding further degradation of the part and possible deterioration of the machine. The example illustrates the relevancy of the proposed approach.

As mentioned at the beginning of the article, the implementation of the ontology is performed in "Protégé" software. The top of Fig. 13 shows the part of the ontology structure that supports the detection process by defining the set of classes, relationships and attributes necessary for knowledge structuring. The attributes enable a description of the operational context. The bottom of Fig. 13 presents the instance of tool failure and its operational context that was detected in the use-case (Fig. 12) and stored in the knowledge base.

Beyond the detections, the traceability of the instances of tool failure detection and their operational context can be reused for diagnosis. They can be analyzed, by the inference engine based on another set of business rules, in order to identify the main faulty cutting tools and contexts, so as to determine the cause of the occurrence and to solve it. These causes are also considered as inferred knowledge.

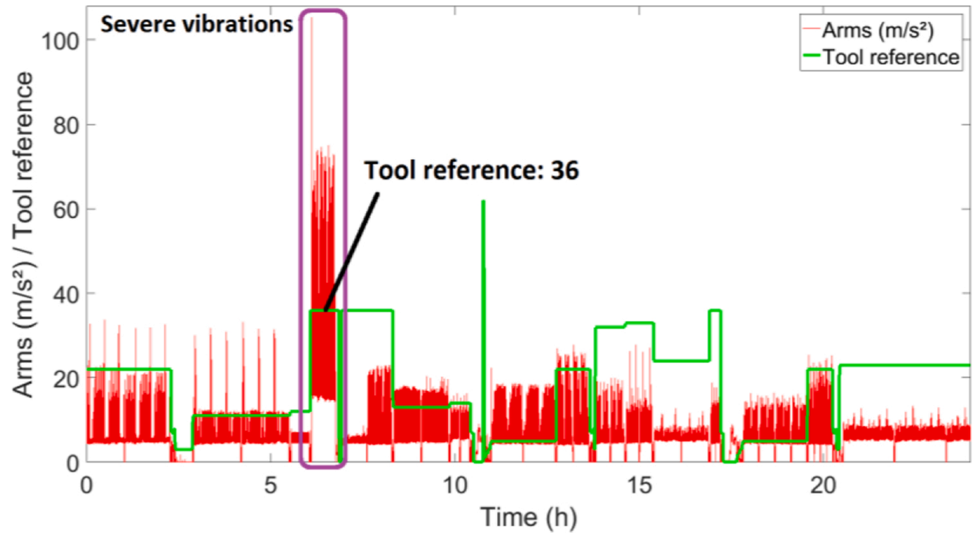


Fig. 11. Raw data of vibration level Arms during the 136th day of production.

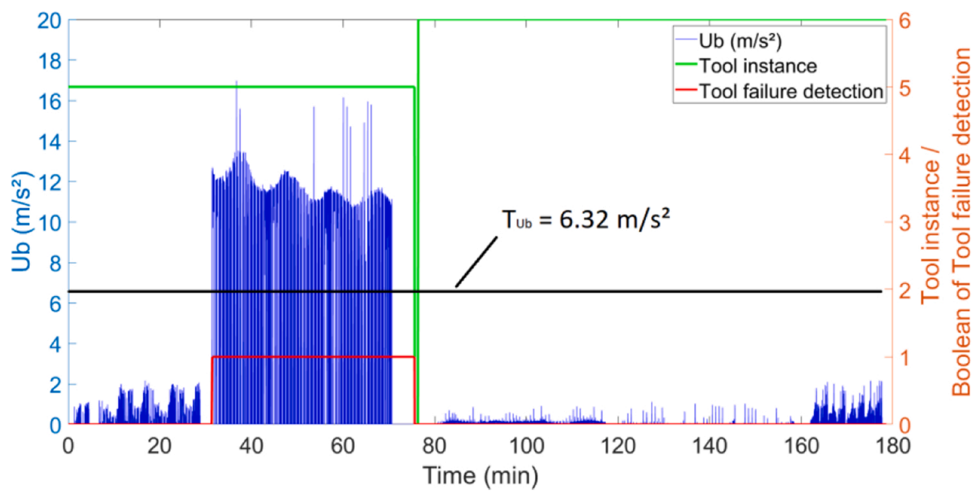


Fig. 12. Visualization of Ub criterion (in blue) for tool reference 36, the index of tool instance (in green) and the Boolean for tool failure detection (in red) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

4.5. Use case findings

This case study proves the feasibility of the proposed framework. From a research point of view, this work allows us to conclude that an efficient design of the Digital Shadow is crucial to improve the performance of manufacturing systems. This objective was achieved in this work by creating an integrated data-knowledge closed loop which combines a data-driven approach with a knowledge-driven approach, where expert rules are extracted and inferred in order to interpret and augment the results of data processing.

The example of "tool failure detection" presented in the case study illustrates the relevancy of the proposed approach in enhancing the decision aid efficiency on the manufacturing shop floor. Direct knowledge (business rules from experts) and derived knowledge (new insights from data processing using the proposed data mining method) enables the immediate detection of tool failure with high accuracy. Consequently, the incident detection enables the machine-tool to stop and avoids further degradation of the part and possible deterioration of the machine. Then, diagnosis and improvement can be performed based on the capitalized instances of tool failure and their operational context. Hence, performance on the shop floor is enhanced.

It is worth noting that, besides the tool failure detection, other incidents are also addressed (e.g. chatter, collision, faulty program, etc.) as described in Section 3.1.1. Furthermore, the proposed knowledge management engine enables the capitalization of positive feedback in the form of possible solutions proposed by the operators/engineers to solve abnormal phenomena. In this way, when similar phenomena are detected, these solutions are directly suggested and reused as best practices.

5. Conclusion

To enhance intelligent manufacturing, data applications based on the Digital Twin paradigm represent key technological solutions. In the presence of big or complex data, the Digital Shadow plays a crucial role in achieving the convergence between physical and virtual spaces. In this paper, a suitable design of Digital Shadow is proposed based on the combination of two principal modules: a knowledge management system and a data analytics system. An ontology-based knowledge model is used to structure the common repository which capitalizes on different types of insights, such as industrial data and knowledge, business rules provided by experts, and knowledge learned from data analytics. This knowledge management method enables the integration of inferred knowledge. The proposed data analytics method relies on unsupervised

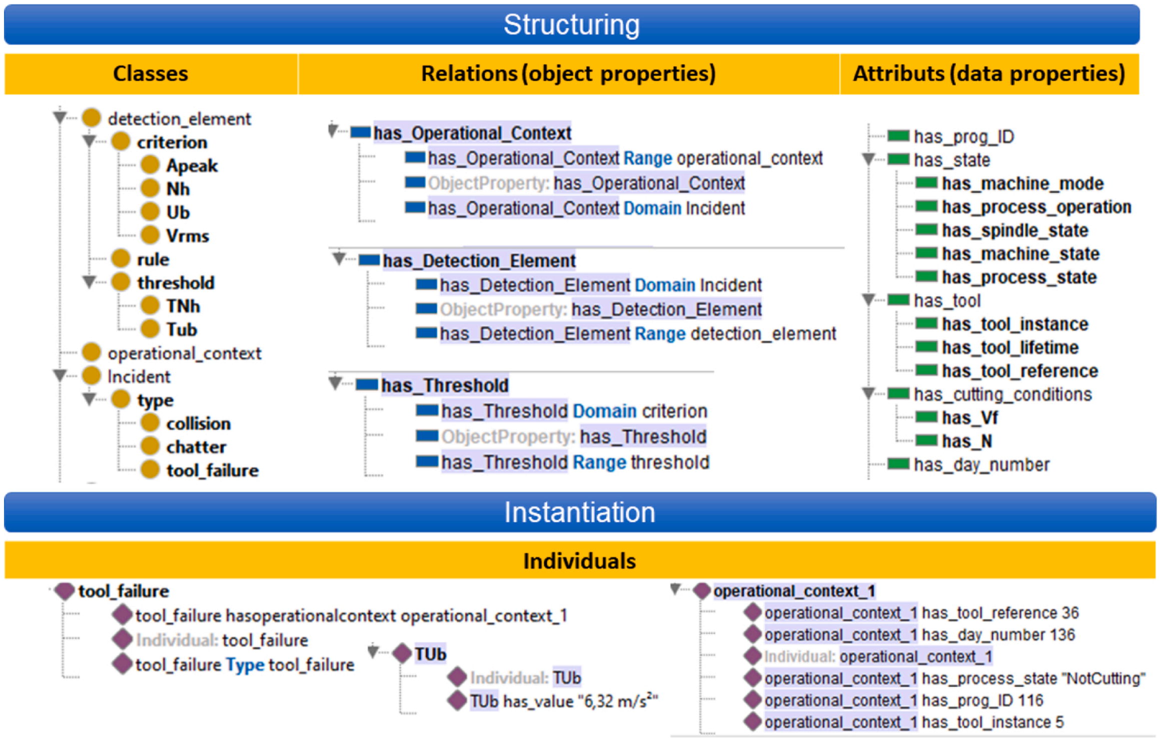


Fig. 13. Structuring and instantiation of the knowledge base in "Protégé" software.

machine learning for the classification of machining contexts and for the detection of detrimental incidents. Reliable monitoring criteria and the deciphered operational context enable the interpretation and understanding of phenomena that occur during the machining process and thus enhance the production performance by supporting decision making. Tool failure detection was presented in the case study, showing the feasibility of the proposed approach. More applications (i.e. collision, chatter, etc.) are under development. Further work will focus on the identification the relationship of causes and effects in order to carry out diagnosis analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was funded by the French National Research Agency (ANR). We gratefully acknowledge the financial support within the ANR SmartEmma project (ANR-16-CE10-0005).

References

[1] Jain S, Shao G, Shin SJ. Manufacturing data analytics using a virtual factory representation. Int J Prod Res 2017;55(18):5450–64.
[2] Brettel M, Friederichsen N, Keller M, Rosenberg M. How virtualization, decentralization and network building change the manufacturing landscape: an industry 4.0 perspective. Int J Mech Ind Sci Eng 2014;8(1):37–44.
[3] Wollschlaeger M, Sauter T, Jasperneite J. The future of industrial communication: automation networks in the era of the internet of things and industry 4.0. IEEE Ind Electron Mag 2017;11(1):17–27.
[4] Uhlemann THJ, Lehmann C, Steinhilper R. The digital twin: realizing the cyber-physical production system for industry 4.0. Procedia CIRP 2017;61:335–40.
[5] Lu Y, Liu C, Kevin I, Wang K, Huang H, Xu X. Digital twin-driven smart manufacturing: connotation, reference model, applications and research issues. Robot Comput Integr Manuf 2020;61:101837.
[6] Schuh G, Kelzenberg C, Wiese J, Ochel T. Data structure of the digital shadow for systematic knowledge management systems in single and small batch production. Procedia CIRP 2019;84:1094–100.
[7] Landherr M, Schneider U, Bauernhansl T. The application center industrie 4.0-industry-driven manufacturing, research and development. Procedia CIRP 2016;57:26–31.
[8] Tao F, Qi Q, Liu A, Kusiak A. Data-driven smart manufacturing. J Manuf Syst 2018;48:157–69.
[9] Monostori L, Kádár B, Bauernhansl T, Kondoh S, Kumara S, Reinhart O, Sauer G, Schuh WSihn, Ueda K. Cyber-physical systems in manufacturing. CIRP Ann Manuf Technol 2016;65(2):621–41.
[10] Qi Q, Tao F, Hu T, Anwer N, Liu A, Wei Y, Nee AYC. Enabling technologies and tools for digital twin. J Manuf Syst 2019. In press.
[11] Glaessgen E, Stargel D. The digital twin paradigm for future NASA and US air force vehicles. 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 2012:1818.
[12] Liu M, Fang S, Dong H, Xu C. Review of digital twin about concepts, technologies, and industrial applications. J Manuf Syst 2020. In press.
[13] Schleich B, Anwer N, Mathieu L, Wartzack S. Shaping the digital twin for design and production engineering. CIRP Annals 2017;66(1):141–4.
[14] ISO DIS23247-1. Automation systems and integration - Digital Twin framework for manufacturing - Part 1: overview and general principles (under development). 2019.
[15] Bao J, Guo D, Li J, Zhang J. The modelling and operations for the digital twin in the context of manufacturing. Enterp Inf Syst 2019;13(4):534–56.
[16] Huang S, Wang G, Yan Y, Fang X. Blockchain-based data management for digital twin of product. J Manuf Syst 2020;54:361–71.
[17] Liu C, Vengayil H, Zhong RY, Xu X. A systematic development method for cyber-physical machine tools. J Manuf Syst 2018;48:13–24.
[18] Cai Y, Starly B, Cohen P, Lee YS. Sensor data and information fusion to construct digital-twins virtual machine tools for cyber-physical manufacturing. Procedia Manuf 2017;10:1031–42.
[19] Choi S, Kang G, Jun C, Lee JY, Han S. Cyber-physical systems: a case study of development for manufacturing industry. Int J Comput Appl Technol 2017;55(4):289–97.
[20] Tong X, Liu Q, Pi S, Xiao Y. Real-time machining data application and service based on IMT digital twin. J Intell Manuf 2019:1–20.
[21] Wang J, Ye L, Gao RX, Li C, Zhang L. Digital Twin for rotating machinery fault diagnosis in smart manufacturing. Int J Prod Res 2019;57(12):3920–34.
[22] Padovano A, Longo F, Nicoletti L, Mirabelli G. A digital twin based service oriented application for a 4.0 knowledge navigation in the smart factory. IFAC-PapersOnLine 2018;51(11):631–6.
[23] Qiao Q, Wang J, Ye L, Gao RX. Digital twin for machining tool condition prediction. Procedia CIRP 2019;81:1388–93.
[24] Liu J, Zhou H, Tian G, Liu X, Jing X. Digital twin-based process reuse and evaluation approach for smart process planning. Int J Adv Manuf Technol 2019;100(5–8):1619–34.

[25] Mourtzis D, Vlachou E. A cloud-based cyber-physical system for adaptive shop-floor scheduling and condition-based maintenance. J Manuf Syst 2018;47:179–98.
[26] Zhuang C, Gong J, Liu J. Digital twin-based assembly data management and process traceability for complex products. J Manuf Syst 2020. In press.
[27] Schützer K, de Andrade Bertazzi J, Sallati C, Anderl R, Zancul E. Contribution to the development of a Digital Twin based on product lifecycle to support the manufacturing process. Procedia CIRP 2019;84:82–7.
[28] Zhang H, Liu Q, Chen X, Zhang D, Leng J. A digital twin-based approach for designing and multi-objective optimization of hollow glass production line. IEEE Access 2017;5:26901–11.
[29] Bauernhansl T, Krüger, Schuh G. WGP-standpunkt industrie 4.0. German Academic Society for Production Engineering; 2016. https://wgp.de/wp-content/uploads/FINAL_WGP_Standpunkt_2025.pdf.
[30] Kritzinger W, Karner M, Traar G, Henjes J, Sihn W. Digital Twin in manufacturing: a categorical literature review and classification. IFAC-PapersOnLine 2018;51(11):1016–22.
[31] Schluse M, Priggemeyer M, Atorf L, Rossmann J. Experimentable digital twins—streamlining simulation-based systems engineering for industry 4.0. IEEE Trans Industr Inform 2018;14(4):1722–31.
[32] Dahmen U, Rossmann J. Experimentable digital twins for a modeling and simulation-based engineering approach. IEEE International Systems Engineering Symposium 2018:1–8.
[33] Schuh G, Jussen P, Harland T. The digital shadow of services: a reference model for comprehensive data collection in MRO services of machine manufacturers. Procedia CIRP 2018;73:271–7.
[34] Jarke M, Schuh G, Brecher C, Brockmann M, Prote JP. Digital shadows in the internet of production. ERCIM News 2018;115:26–8.
[35] Riesener M, Schuh G, Dölle C, Tönnes C. The digital shadow as enabler for data analytics in product life cycle management. Procedia CIRP 2019;80:729–34.
[36] Negri E, Berardi S, Fumagalli L, Macchi M. MES-integrated digital twin frameworks. J Manuf Syst 2020;56:58–71.
[37] Stecken J, Ebel M, Bartelt M, Poeppelbuss J, Kuhlenkötter B. Digital shadow platform as an innovative business model. Procedia CIRP 2019;83:204–9.
[38] Legnani E, Ierace S, Cavalieri S. Towards a reference model for after-sales service processes. Adv Prod Manage Syst 2007:289–96.
[39] Noy N, McGuiness D. Ontology Development 101: A Guide to Creating Your First Ontology. Standford University; 2001.
[40] Fankam C. OntoDB2 : un système flexible et efficient de Base de Données à Base Ontologique pour le Web Sémantique et les données techniques. PhD thesis of ENSMA. France: Futuroscope; 2009.
[41] Krima S, Barbau R, Fiorentini X, Foufou S, Sriram R. OntoSTEP : OWL-DL ontology for STEP. In: International Conference of Product Lifecycle Management; 2009.
[42] Maleki E, Belkadi F, Ritou M, Bernard A. A tailored ontology supporting sensor implementation for the maintenance of industrial machines. Sensors 2017;17(9):2063.
[43] Godreau V, Ritou M, Chové E, Furet B, Dumur D. Continuous improvement of HSM process by data mining. J Intell Manuf 2019;30(7):2781–8.
[44] Lenz J, Wuest T, Westkämper E. Holistic approach to machine tool data analytics. J Manuf Syst 2018;48:180–91.
[45] Ritou M, Belkadi F, Yahouni Z, Da Cunha C, Laroche F, Furet B. Knowledge-based multi-level aggregation for decision aid in the machining industry. CIRP Annals 2019;68(1):475–8.
[46] Linger H, Fisher J, Wojtkowski WG, Wojtkowski W, Zupancic J, Vigo K, Arnold J. Constructing the Infrastructure for the Knowledge Economy: Methods and Tools, Theory and Practice. Springer Science & Business Media; 2013.
[47] De Castelbajac C, Ritou M, Laporte S, Furet B. Monitoring of distributed defects on HSM spindle bearings. Appl Acoust 2014;77:159–68.
[48] Maugis C, Celeux G, Martin-Magniette ML. Variable selection for clustering with Gaussian mixture models. Biometrics 2009;65(3):701–9.
[49] Wang Z, Ritou M, Da-Cunha C, Furet B. Contextual classification for smart machining based on unsupervised machine learning by Gaussian mixture model. Int J Comput Integr Manuf 2020. https://doi.org/10.1080/0951192X.2020.1775302. in press.
[50] Whittaker J. Graphical models in applied multivariate statistics. Wiley Publishing; 2009.
[51] Wirth R, Hipp J. CRISP-DM: towards a standard process model for data mining. 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining 2000:29–39.
[52] Bahetil R, Gill H. Cyber-physical Systems. In: Samad T, Annaswamy AM, editors. The Impact of Control Technology. IEEE Control Systems Society; 2011. p. 161–6. In this issue.
[53] Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, et al. CRISP-DM 1.0: Step-by-step data mining guide. SPSS inc 2000;9:13.
[54] Fayyad U, Piatetsky G, Smyth P. From data mining to knowledge discovery in databases. AI Magazine 1996;17(3):37–53.