Doctoral Dissertation
Doctoral Program in Management, Production and Design (34^{th}cycle)

Hybrid modeling to support the smart manufacturing

concepts, theoretic contributions and real-case applications about Hybrid and Wisdom-based Systems

By

Emiliano Traini

******

Supervisor(s):
Prof. Franco Lombardi, Supervisor

Doctoral Examination Committee:
Prof. Frederic Segonds, Ecole Nationale Superieure d'Arts et Metiers
Prof. Martì Casadesus Fa, Universitat de Girona
Prof. Marco Ghirardi, Politecnico di Torino
Prof. Roberto Fontana, Politecnico di Torino
Adjunct Prof. Valentino Razza, Istituto Italiano di Tecnologia

Politecnico di Torino
2022

Contents

Chapter 1

Introduction

Just as the microscope empowered our naked eyes to see cells, microbes, and viruses, thereby advancing the progress of biology and medicine; or as the telescope opened our minds to the immensity of the cosmos and has enabled humankind to explore and understand the space in spectacular detail, in the current century, ComputingScope (computational "instruments" for "viewing" and analyzing data) will help successfully decipher and navigate another infinite: the staggeringly complex multi-modal (text, image, video, and sensors) information in all facets of our lives. [1]

This work has been written in the context of widespread use of technologies and paradigms of the Industry 4.0, or also initially called smart manufacturing. Answering to the need of a unique value standard for the company, valid from shareholders to stakeholders, the same context sees the birth of the Industry 5.0 idea.

1.1 Author's background

This manuscript is the result of the doctoral program in "Management, Production and Design" that I approached with a master's degree in mathematical engineering obtained from the same university, the Polytechnic University of Turin. During these years of doctoral studies and research grants, the topics addressed and with published scientific results are as follows:

- Machine Learning (ML) methods and hybrid forecasting systems for Predictive Maintenance, [2] [3] [4] [5]

- analytical models and simulation methods for cycle time, throughput and energy consumption optimization for automated warehouse systems, [6] [7] [8] [9]

- functional integration between Manufacturing Execution Systems and Enterprise Resource Planning (ERP) and Product Lifecycle Management (PLM) systems toward a new factory knowledge, [10] [11] [12] [13] [14]

- design and simulation of a Battery Swapping System (BSS) for Electric Vehicles (EVs). [15]

During the same years, I have been involved in the following research projects.

- **Safe&Green Intralogistic System with 4.0 integrations (SaGrIS4.0)** belonging to the MESAP polo. The aim of the project was to integrate the innovative ESMARTSHUTTLE®SAFE prototype with the high-efficiency automated warehouse within a manufacturing plant in order to allow real-time management of all data collected in the plant and to increase its overall efficiency from both an economic and energy perspective. From an economic point of view, the integration will both avoid costly downtime due to inefficient warehouse management and reduce product picking and shipping times by optimizing warehouse utilization based on production data. From an energy perspective, the use of the ESMARTSHUTTLE®system, which requires much lower energy consumption than conventional handling systems, will achieve significant energy efficiency. The innovative automated warehouse integrated into the production plant has been implemented in a pasta production and storage system.

- **Hierarchical Open Manufacturing Europe (HOME)**, an European project searching for solutions for environments where the aggregate information is there, but not where it is needed, when it is needed, who needs it. This makes it extremely complex to ensure the economic, social, environmental, and energy sustainability of manufacturing in sectors that are centuries old (automobile) or more (textile). The overall goals of the project are seemingly simple: to make manufacturing lean, smart, aware, and sustainable. The project places at the center of any manufacturing conversion process the human being, who is the entity that determines its success or failure. That is why the name of the project: Hierarchical, because it does not refer to social organization, but to information systems architecture, as an alternative to anarchic architectures, in which each sector of manufacturing has its own database-administrative, commercial, productive-not communicating with each other; Open, to free the shop floor from the "Babel" of proprietary protocols that prevent the transfer of innovation; and Manufacturing Europe, because the project aspires to be a model for the entire manufacturing industry, not only Italian, but European, because for the first time we can consider ourselves a single market of evolved consumers, with over 500 million people.

- **CAPTure aNd foStEr additive manufacturing knowlEdge for luxury industry (CAPT'N'SEE)**, co-funded by the European Union, is a training program dedicated to professionals willing to enhance their expertise in the use of Additive Manufacturing (AM) technologies. The project is particularly addressed two core steps of the AM value chain that have been neglected so far: the AM early design stages, and the Manufacturing Execution Systems (MES) stage allowing real-time control of processed parts.

1.2 Motivations of the research

The main motivation of the research is the desire to investigate hybrid systems and their application in the manufacturing sector. Taking the different research projects discussed as examples, different types of models were used depending on the data available, the complexity of the system, and the type of analysis to be performed. It is also true that in many case studies different models have been applied in order to obtain the same estimates of variables values, so much so that comparisons between such models have been used as a technique for verifying the developed model: for example, discrete-event simulation methods have been used to verify the analytical models used to estimate cycle time, throughput, and energy consumption of average handling and its standard deviation for automated warehouse systems, or different techniques have been used to map and describe the resources of an enterprise information system according to the level of detail needed. These examples are just a few of the various research works in the literature that use an approach that can be considered a hybrid approach to modeling.

This thesis work aims to propose a method or rather a framework for designing an information system that not only includes, but integrates different modeling techniques in order to optimize the stages of mapping the information flows of the system and describing the knowledge needed in order to achieve the awareness that can guide toward optimal choices. Such hybrid systems, i.e., hybrid model systems, turn out to be an interesting research topic as they are widely used even though the scientific literature provides neither a clear definition and comprehensive results regarding their benefits over single models and under different hybridization choices. Another important open research question specifically concerns system knowledge and, in particular, how to integrate so-called prior knowledge and how to structure a framework that provides for the addition of new knowledge sources over time. [1]

The main scientific contribution that this work seeks to bring, therefore, is a proposal for a design method that promotes to the reader, or accentuates, hybrid thinking, that is, designing an information system by considering different models, separately and simultaneously, in order to obtain more reliable descriptions and predictions of the state of the system, ensuring greater resilience of the system as it is able to exploit the strengths of different prediction models. This certainly pushes research efforts toward the concepts of Enterprise Information Systems (EISs) and Knowledge-Based Systems (KBSs), and thus toward the study of how to use different information mapping techniques and different approaches of variables modeling (the main ones being data-driven and those based on laws of physics, analyzing how integrate human-driven ones, i.e., variables models based completely on manufacturing human know-how).

Hybrid modeling, considering different modeling techniques, aim to avoid the unscrupulous use of one single family of models that are often non-sustainable. It is unequivocal that the sustainable manufacturing needs an appropriate digital information system that has (or that is design with) an awareness of the enterprise's objectives and the impact of its use by the enterprise's resources. Today it is required that this awareness is increasingly comprehensive and effective, in other words, that it follows the 5.0 vision by making use (or being able to make use) of all 4.0 technologies. In order to guarantee the sustainability, an hybrid model has to consider the impact due to energy consumption or hardware production and installation, but often the non-sustainability is also due to maintenance costs or the inability of humans to use the system.

Since it is necessary to introduce the concept of sustainability in the choice of models to be used, it is evident that it is not enough to think exclusively in terms of mapping information flows and cognitive processes (based and therefore generating knowledge), but it is necessary that a Cyber-Physical System (CPS) possess the ability to finalize these cognitive processes to increase the effectiveness and add the value of business processes. This ability requires a function that can be associated with the human concept of judgment: both the judgment of society, or how a subject relates to global objectives, and judgment towards oneself, or how the same subject judges itself on the basis of its own personal objectives. Considering the need of such smart CPS, from 2020, the Commission department of Research and Innovation has published several documents on the European vision of Industry 5.0, i.e., the era where industries are focused on challenges for society, including the preservation of resources, climate change and social stability. In this sense, the subject's wisdom could be described by Industry 5.0's own points, that, nowadays, according the European view, can be resumed by the pillars of human-centrality, sustainability and resilience. The Figure 1.1 shows few results obtain on the Scopus platform related to the concept of smart manufacturing and Industry 5.0.

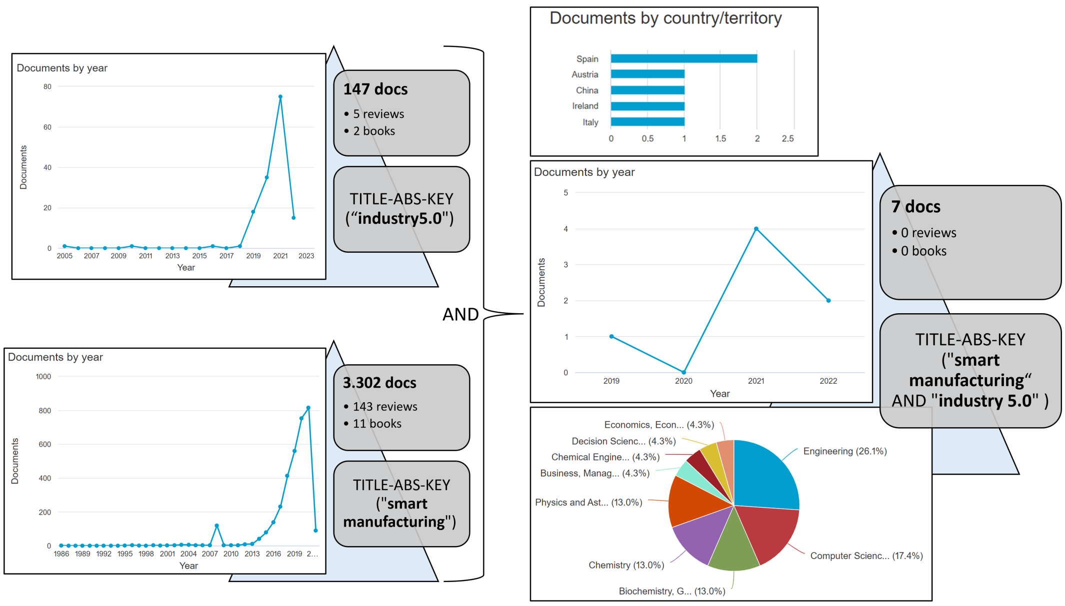

Fig. 1.1 Scopus research: results related to the individual and intersecting concepts of Industry 5.0 and smart manufacturing.

1.3 Aim of the research

The aim of this work is contributing to the research concerning concepts and techniques used to manage the enterprise digital information, i.e., to power it by producing or making available data and to use that information to create a manufacturing enterprise that can be considered sustainable in this era: the dawn of the 5.0 industrial philosophy. This work concerns the formalization and application of a framework based on hybrid models and with the aim of creating digital twins [16] of products and manufacturing processes, to provide them with explicit knowledge. The work assumes that the cooperation between formal models of explicit knowledge and data driven algorithms could contribute to the efficacy and efficiency in cognitive processes, as a whole. In particular, this framework is design to be focused on manufacturing processes and products and on (i) effectively managing the history of data relating to them, (ii) comprehensively describing the current condition of these entities, and (iii) accurately estimating future states, in order to support the manufacturing decision-making activity.

In detail, this thesis is a work dealing with (i) awareness Knowledge-Based Systems (KBSs) and (ii) hybrid systems as base concepts for designing a digital platform aiming of supporting a specific physical manufacturing environment. The model proposed by this work is based on these two concepts and, basically, it is a theoretical formalization of a digital platform belonging to a Cyber-Physical System (CPS) that uses hybrid models in order to achieve a 5.0 wisdom, i.e., in order to promote and follow a 5.0 awareness as digital component of information management for a generic 4.0 manufacturing system. The hierarchical structure of the 4 levels of Data, Information, Knowledge and Wisdom (DIKW) is proposed as a method capable of (i) making use of data-lakes, information flows and knowledge processes to make conscious decisions, with wisdom for the precisely, and (ii) developing the concept of hybrid system by characterizing the 4 hybrid subsystems: hybrid data sources, DB and computer networks hybrid models, hybrid models for estimating state variables, and hybrid decision support methods.

Case-study applications are presented in order to test the robustness of the proposed framework in terms of generality and specificity in the manufacturing sector. Another expected outcomes related to the proposed framework is its evident contribution towards a 5.0 smart manufacturing. The main contribution of the case study is investigating on the synergy between Physics-Based (PB) models

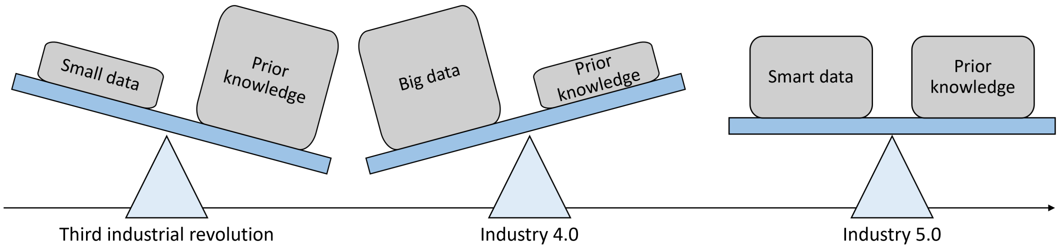

Fig. 1.2 Representation of the weight put on types of information in different modeling paradigms inspired on a review work about hybrid modeling and Industry 4.0 (the Industry 5.0 scenario is considered enabled by hybrid modeling techniques of recent years). [17]

and Machine Learning (ML) or completely data-driven models. Since PB models require few data only to calibrate the effectiveness of the understanding of the system dynamics, and ML models are completely data-based, the aim is to develop a hybrid model that can work with significant performance already at the beginning of the data acquisition activity, thanks to the physics approach, and improve as the data increasing, thanks ML to algorithms. The equilibrium shown by the Figure 1.2 is hypothesized mainly based on this synergy.

1.4 Research questions

1.4.1 RQ1: concerning Industry 5.0

Is the DIKW-schema able to support the design of a Decision Support System (DSS) integrated in a 5.0 smart manufacturing context?

Actually, the size of the scientific literature regarding the current Industry 5.0 does not justify a full interest on the subject. In any case, the concepts of sustainability, resilience and human-centrality are clear pillars from which to build a variables system on which decision-making processes are based to achieve a system able of managing changes in the factory and in the system in which such factory operates, thanks to the collaboration with the human resources of the factory, and in order to make the manufacturing CPS sustainable. That literature regarding the Industry 5.0 underlines the need to find a way to incorporate into the design of AI services some sort of overall vision of the enterprise that can guide AI decisions with factors that include the welfare of broader systems, even broader than the enterprise system itself.

In this sense, the DIKW pyramid is a proposal along these line, where the wisdom subsystem represents (i) a concept that drives designers to include such universal factors as the target of the service to be developed, and (ii) a general component common to all the services in the factory in order to create a common vision followed by all digital agents, and non-digital agents, who are able to make effective decisions within the factory.

In order to answer this question, this work contains the description of the following research activities:

- literature review on the DIKW schema,
- literature review on the Industry 5.0,
- framework design towards 5.0 view trough the DIKW principle,
- case study on a Total Productive Maintenance (TPM) system in order to test the applicability of a DIKW agent-based model.

1.4.2 RQ2: concerning hybrid systems

What can be a formal and comprehensive definition of hybrid systems employed as manufacturing decision support components?

The second research question refers primarily to the definition of hybrid system and both the analysis of the literature and the proposed framework contribute for the answer to such question: in the literature the use of this term is employed in different ways and occasionally even in conflicting ways. Continuing, the second research question also refers to how to measure the contribution of a hybrid system, which are the models to use in such system and which is its structure, i.e., how these models are integrated.

In order to answer this question, this work contains the description of the following research activities:

- literature review on hybrid systems,
- discussion of the hybridization referred to data bases, information systems and Knowledge-Based Systems (KBSs) in order to provide a more comprehensive definition,

- analysis of the human cross-disciplinarity characteristic as a fundamental component of each digital components involved in a Decision-Support System (DSS),
- case study focused on physics-based and (sensors) data-driven hybrid model for KBSs.

1.4.3 RQ3: concerning (human) manufacturing prior knowledge

The third research question reads as follows:

Is hybrid modeling a paradigm that standardize the use of human prior knowledge in manufacturing decision support systems? If so, how?

The third question investigates how to integrate human prior knowledge in the framework and how to structure such framework in order that it is able to generate new knowledge and to receive knowledge from new sources over time.

In order to answer this question, this work contains the description of the following research activities:

- literature review on hybrid systems,
- discussion of the hybridization referred to Knowledge-Based Systems (KBSs),
- analysis of the human cross-disciplinarity characteristic as a fundamental component of each digital components involved in a Decision-Support System (DSS),
- theoretical proposition of the human prior knowledge involved with simulation, physics-based and (sensors) data-driven models in hybrid KBSs.

1.5 Thesis outline

The thesis consists of the following five chapters.

1. **Introduction** The first chapter has introduced the background of the author, the motivations and the aim of the research and the research question.

2. **Manufacturing context.** In the second chapter there is a quick introduction to the main manufacturing concepts on which this thesis is based: lean manufacturing from which is important to understand variables and methodologies to eliminate wastes, Enterprise Information Systems (EISs) focusing on ERP, PLM, MES and their integration, Industry 4.0 and its technologies, and finally the Industry 5.0 and the European manufacturing view.

3. **Design framework for smart data-driven manufacturing services.** The chapter is regarding the description of the proposed hybrid wisdom-based framework. It includes an analysis of manufacturing processes and the opportunities they offer for the proposed framework. After giving basic notions about agent-based system, DIKW-structures and hybrid modeling, the element of the framework, the agent, is presented with DIKW levels, the hybrid structure of which is subsequently discussed level by level. Finally, a proposal of framework evaluation metrics and the expected impacts in Industry 4.0 and Industry 5.0 is provided.

4. **HW-TPM system for CNC machine tools.** The fourth chapter is focused on the application of the framework to design a Tool Condition Monitoring (TCM) system for a Total Productive Maintenance (TPM). Results obtain with a real case application are provided on milling process: monitoring and optimize the changeover of the milling cutters using open data for results replication.

5. **Conclusions.** The last chapter (i) provides the summary of the work, (ii) gives conclusive remarks, (iii) discusses the research questions by discussing the results from theory and case studies also from a holistic point of view, and (iv) gives guidelines for future improvements for the proposed frameworks and about new case studies.

Chapter 2

Manufacturing context

The aim of this chapter is to clarify definitions, standards, and theory used in the whole treatment. This treatment is specific in the manufacturing context where the affirmation of 4.0 paradigms, Information Technologies (ITs) and where the 5.0 philosophy is emerging.

Smart manufacturing can be defined as the extensive application of computer-integrated manufacturing and advanced intelligence systems to enable rapid manufacturing of new products, dynamic response to product demand, and real-time optimization of manufacturing production and supply-chain networks. [18] Smart manufacturing is considered as the evolution of Intelligent Manufacturing (IM), where IM being knowledge-based, whereas smart manufacturing is data-driven and knowledge-enabled. Smart manufacturing uses Artificial Intelligence (AI) techniques to learn directly from data and assist decision making, in contrasts with the "expert system" approach that aims to mimic the rules from human experts with the help of analysts who translate human rules and context expertise into software models. [19] [20]

The term smart manufacturing is thus a concept peculiar to this fourth industrial revolution and it sets the goal for process and IT platform designers to achieve. In the next paragraphs, in addition to dealing with this concept narrowly related to Industry 4.0 and the Cyber-Physical System (CPS), the lean philosophy is first introduced which is a manufacturing philosophy that chronologically anticipates this revolution but today represents a paradigm synergistic to Industry 4.0 as shown by several works in the literature. Finally, in that chapter, the very young concept of the fifth industrial revolution is also presented. The industry 5.0 is the result of the volunteer of scientists and industrialists to decide the best general vision or approach for the use of 4.0 technologies, paying attention to the issues of human-centric, sustainability, and resilience.

2.1 Lean manufacturing

The term lean was designed in 1992 by researchers of MIT Womack and Jones, in their best-seller book "The Machine that Changed the World", [21] outlining the system of production that allowed the Japanese company Toyota's results clearly superior to all competitors in the world Since then, thousands of excellent organizations in the world have adopted the lean model, as in the services industry, as applicable to all operational processes, not only strictly productive, but also logistical, administrative, or product design and development.

Lean is a set of principles, methods, and techniques for the management of operational processes, which aims to increase the value perceived by the end user and to systematically reduction of the waste. This is only possible with the involvement of people committed to continuous improvement. The goal of Lean Production is "do more with less and less": less time, less space, less effort, fewer machines, and less material.

2.1.1 Lean principles and 3MUs

Lean Thinking emphasizes how the lean, as well as a method to be applied, is first and foremost a mindset, a way of thinking that inspires the same method. Lean is based on five principles:

1. **value**, i.e., the starting point is always the definition of the value from the perspective of the customer. Value is only what the customer is willing to pay, all the rest is waste, and should be deleted,

2. **mapping**, i.e., o eliminate waste must "map" the value stream, which outline all the activities that make up the operating process distinguishing between those value-added and non-value-added,

3. **flow**, i.e., the process of value creation is seen as a flow, which must slide in a continuous manner, with the relative reduction of the throughput time (lead time) of the material,

4. **production "pulled"**, i.e., customer satisfaction means producing only what he wants, when he wants and only what he wants. The production is so "stretched" by the customer, rather than "push" from the producer,

5. **perfection**, i.e., perfection is the benchmark to which you must strive endlessly through continuous improvement, and corresponds to the complete elimination of waste.

3MUs (**muda**, **mura** and **muri**) is a lean management tool designed to cut waste, and improve processes and work flow:

- **muda**, i.e., any activity that consumes resources (including time) but creates no value for a customer,
- **mura**, i.e., variation in the operation of a process not caused by the end customer,
- **muri**, i.e., overburden on equipment, facilities, and people caused by Mura and Muda.

Given some of the major reasons project fail are planning and execution driven. Therefore, 3MUs diagnostic tool can be used, both, at planning and execution stages to improve the quality of project performance. We explain this below with some possible uses. However, following is neither an exhaustive nor a perfect list of items that can be done using the philosophy of 3MUs. Hence, the list should be taken as an illustration for explanation purposes.

Using muda principle will mean not to do the following:

- assigning work without matching relevant experience and skills to the task needs,
- designing new tools and templates rather than looking for possibilities of using the available tools,
- assigning more resources than the effort requirements or work needs,

- assigning requirements gathering work to someone with no knowledge and prior experience of similar projects,
- asking a new team member to work on creating Work Breakdown Structure, activity definition and sequencing, and resource estimations,
- assign someone with no experience and knowledge of risk planning and management,
- taking a tick box approach to assigning resources to roles,
- only involving few selected team members in risk and quality management and not involving the entire team.

Using mura principle, the aim will be to avoid discrepancies, interruptions and irregularities in work flow. It could mean doing several things such as ensuring:

- setting project norms and behaviors requirements at the start of the project, and diligently making sure that all team members know them and abide by them,
- having clearly delineated reporting relationships, roles and responsibilities,
- using project dashboard and charts to monitor work on regular basis,
- ensuring digital systems, and other resources at the physical location of project work are smoothly functioning,
- having work flow charts displayed at a prominent location for clarity of understanding,
- having troubleshooting and escalating procedures worked out and disseminated among the team members for their use in case of any potential situations,
- having back-up plan for key resource attrition to avoid disruptions to workflow if anyone of them leave the project mid-way,
- having clear Human resource policies that delineate rewards and consequences procedures,
- having a balanced workloads for team members,

- have clear procedures for resolving complaints and concerns,
- ensuring stable decision making rather than making decisions on fly and rolling them back latter
- having leaders that command respect,
- ensuring motivation and commitment of team members is maintained,
- ensuring that any changes to procedures should be well considered, consulted and implemented by involving and informing all concerned.

Using muri principle, the aim will be to avoid development of situations that cause stress to team members and process flows. Muri, typically, is an outcome when things are not done as per the principles of muda and mura

Looking at the above list of items for how 3MUs can be potentially used in management and decision making aspects of project management, one gets the feeling that most of these items can be implemented with least documentations and purely active involvement and engagement of team members.Given the leanness of documentation and lightweight processing, it seems that 3MUs is **a tool that allow Small (or better Smart) efforts for potentially Big benefits**.

The key elements of lean production can be represented in the so-called "House of Lean" shown by the Figure 2.1. The four pillars, which will be described later, are: (i) Just-In-Time (JIT), (ii) autonomation (jidoka), (iii) Total Productive Maintenance (TPM or productive maintenance), and (iv) Workplace Organization (WO).

At the base of the pillars there are two fundamental concepts: (i) standardization (standard work), which makes extensive use of visual management, and (ii) continuous improvement (kaizen), which relies on technical specifications of problem solving.

It is important to emphasize that the goal of lean production is rigorously and systematically strive for total cancellation of waste ("zero-vision"), not to its simple reduction. Each pillar has its own target zero: zero inventory for JIT, zero defects for jidoka, zero changeover for TPM, and zero inefficiencies for WO. These lenses, which help to achieve zero waste (muda), they become perceived value to the customer in terms of quality, cost and time.

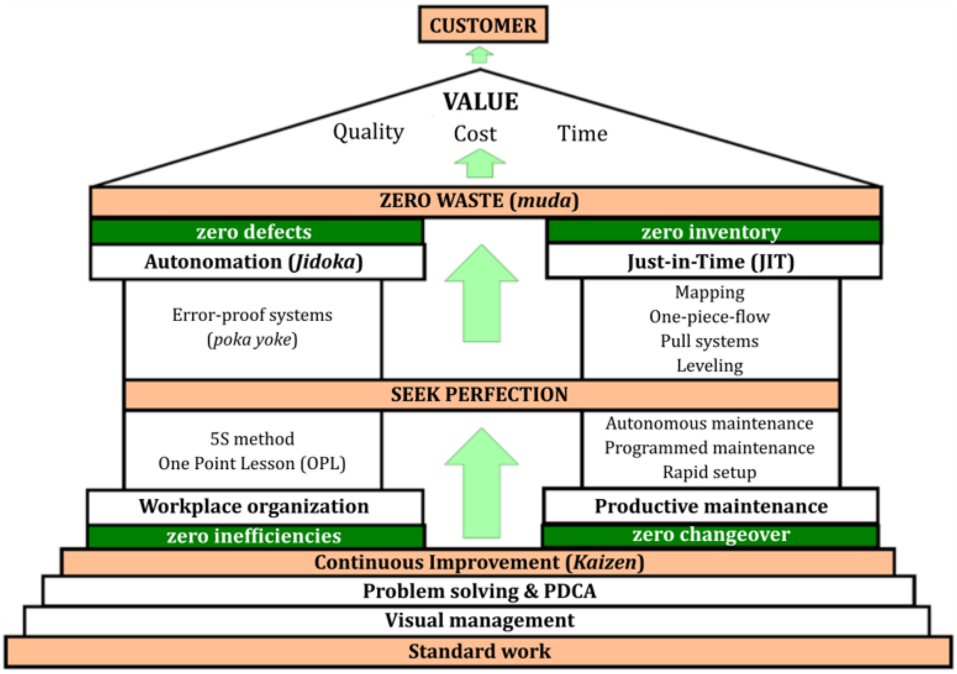

Fig. 2.1 The "Lean House".

2.1.2 Seven wastes (muda)

It is wasteful everything that consumes resources, in terms of cost and time, but without creating value for the customer. In Japanese culture, the concept of waste (muda) has a meaning similar to the western ethic of sin, and is therefore a strong motivation to avoid it. They are further classified into seven types:

1. **transportation**, i.e., when you move resources (materials), and the movement does not add value to the product,

2. **inventory**, i.e., cases where companies overstock themselves in order to manage unexpected demand, production delays, quality, or other problems, however, this does not meet customer's needs and don't add value, increasing storage and depreciation costs,

3. **motion**, i.e., movements of resources (employees or machinery) and products,

4. **waiting**, i.e., whenever goods or tasks are not moving, the "waiting waste" occurs,

	Lean methods			
Waste type	JIT	JIDOKA	TPM	WO
Overproduction	⋆⋆⋆	⋆	⋆	
Defects	⋆	⋆⋆⋆	⋆	⋆
Inventory	⋆⋆⋆	⋆		
Motion	⋆		⋆⋆	⋆⋆⋆
Transportation	⋆⋆⋆			
Over-processing	⋆		⋆⋆	⋆⋆⋆
Waiting	⋆⋆⋆		⋆⋆⋆	

Table 2.1 Waste types and lean method to remove them. [22] [23]

5. **overproduction**, i.e., producing more means that you exceed customer's demand, which leads to additional costs,
6. **over-processing**, i.e., on doing work that does not bring additional value, or it brings more value than required,
7. **defects**, i.e., phenomena that can cause rework, or even worse, they can lead to scrap

Lean aims to eliminate waste through four main methods, which are described in the Table 2.1.

2.1.3 Just-In-Time (JIT) management and 5s methodology

The Just-In-Time (JIT) is a logistics and production method whose aim is to produce and deliver to the customer: (i) only thing required, (ii) only when required, and (iii) only what is required. Together with the autonomation, the JIT is the main pillar of lean production, as it gives speed and flexibility to the system logistics and production and results in progressive reduction of all types of waste as shown by the Table 2.1. In particular, with the Just-In-Time are obtained remarkable reductions of (i) throughput time (lead time) and (ii) space of establishment, needed to contain the flow of production and inventories, due to the reduction of waste from overproduction, unnecessary escort and transportation.

The basic rules of JIT are:

- do not produce if the customer does not require,

- level the question,
- connect all processes to customer demand with simple visual tools (**kanban**).

The proper functioning of the JIT strongly depends on the simultaneous application of all the principles, methods and techniques lean, as this gives the necessary stability to the system. The main operational elements of a Just-In-Time are:

- **continuous flow**,
- the production "pulled" by the customer (**pull system**),
- leveling of production (**heijunka**).

Lean manufacturing or "continuous flow" provides for the progressive reduction of the lot size, tending ideally equivalence "1 lot = 1 unit", i.e., the production and movement of a piece at a time (**one-piece-flow**). In this way, the production flows continuously, without interruption, expectations and semi-finished goods warehouses. The approach of the work stations also reduces wastage of handling. The pull system is a method for controlling the flow of materials based on systematic replenishing only of what is actually consumed. The production is "pulled" (pull) by the client, in the sense that each processing step occurs only if requested by the step of downstream processing. Moreover, the manufacturing processes are "leveled" as there is a constant distribution of workloads between successive stations. In order to make more regulate the activity of a production line, it is necessary to (i) to regulate the demand of the customer (if possible), through leveling deliveries (level selling), and (ii) to regulate the production, through the leveling of the same (heijunka), obtained with the frequent production of small batches of each product alternating with small lots of other products. This will make it easier (i) adapting quickly to changes in demand and (ii) reducing excessive loads.

5S is a workplace organization method that uses a list of five Japanese words: seiri, seiton, seiso, seiketsu, and shitsuke. These have been translated as "sort", "set in order", "shine", "standardize", and "sustain". The list describes how to organize a work space for efficiency and effectiveness by identifying and storing the items used, maintaining the area and items, and sustaining the new organizational system. The decision-making process usually comes from a dialogue about standardization, which builds understanding among employees of how they should do the work.

Japanese name	English 5s	English 5C	Features
Seiri	Sort	Clear	Sort out unnecessary items from the workplace and discard them.
Seiton	Set in order	Configure	Arrange necessary items in good order so that they can be easily picked up for use.
Seisio	Shine	Clean and check	Clean the workplace to make it free from dust, dirt and clutter.
Seiketsu	Standardize	Conformity	Maintain high standard of house keeping and workplace organization.
Shitsuke	Sustain	Custom and practice	Train and motivate people to follow good housekeeping disciplines in autonomous way.

Table 2.2 Key activities for effective 5S implementation at the workplace.

In some applications, 5S has become 6S, the sixth element being safety (safe). [24] Generally, five 5S phases are considered and they are explain Table 2.2.

2.1.4 Autonomation (jidoka)

The autonomation (Jidoka), also described as "intelligent automation" [25] or "automation with a human," is a preventive method of quality management, in which operators and machines will self-activate to identify abnormal processes, understanding the causes and eliminate them quickly. The effectiveness of autonomation from the ability to detect errors before they become defective product to stop the process if necessary, and to make visible the causes of the problems as soon as they occur, thus facilitating the deletion. The main typologies of errors are faulty parts assembling or processing, incorrect positioning, wrong quantities, materials lack, assembling omission, tool settings, wrong identification, slowness, and lack of supervision.

In lean production autonomation the importance significantly higher automation because only the systematic elimination of anomalies can allow continuous flow pulled by the customer's own JIT, as well as making possible large increases in productivity processing multi-machine and multi-process. The autonomation aims to achieve zero defects, then 100% quality, as no customer is willing to tolerate a defective product. It goes well beyond the traditional statistical approach to quality, which is limited to reduce defects by a percentage "acceptable", but does not aim to eliminate them altogether. In industrial production, accept a defect of 0.1% (one per thousand) is to accept a dangerous landing per day at an international airport.

The principle tool of autonomation is the `poka-yoke` that means fail-safe". The Poka-Yoke (P-Y) is a set of devices, mechanisms, or simple expedients designed to prevent errors from becoming defects. They are based on a logic of defect prevention and quality management at the source. They are most effective when they allow absolute control, provide immediate feedback, they are simple, rugged, reliable, economical, and when they require special attention by the operator. It is preferable that such solutions are designed already in the development phase of the product and the process. The fault-tolerant systems allow management of errors and defects in a very short time ("short cycle") than the traditional quality control ("long cycle"). The poka-yoke are classified according to several criteria:

- criterion of effectiveness, i.e., (i) PY of error prevention and control at the source prevents the error, (ii) PY fault prevention prevents the error becomes defect, and (iii) PY detection detects when the defect has already occurred, but allows you to limit the number and, in any case, do not ever get to the customer,
- use policy, i.e., (i) P-Y control manages production downtime and (ii) P-Y warning alerts the operator with sounds or lights,
- criterion of the problem, i.e., (i) P-Y product detects anomalies in the product and (ii) P-Y process detects process abnormalities,
- criterion of operation, i.e., (i) P-Y contact works by contact with the product, (ii) P-Y for fixed values verifies compliance with a standard number of parts or events, and (iii) P-Y for sequence of operations (motion-step) checks the correct repetition of an operation.

In addition to the numerous possibilities offered by modern sensor technology, creativity, technical and organizational design allows an infinite number of simple solutions poka-yoke, suitable for both production processes for operational risks.

2.1.5 Total Productive Maintenance (TPM)

Equipment management has gone through many phases. The progress of maintenance concepts over the years is explained below:

- **Breakdown Maintenance (BM)** refers to the maintenance strategy where repair is done after the equipment failure o stoppage,
- **Preventive Maintenance (PM)** is a kind of physical check up the equipment to prevent equipment breakdown,
- **Predictive Maintenance (PdM)** is often referred to as Condition Based Maintenance (CBM) where maintenance is initiated in response to a specific equipment condition or performance deterioration,
- **Corrective Maintenance (CM)** is the concept to prevent equipment failures is further expanded to be applied to the improvement of equipment so that

the equipment failure can be eliminated (improving the reliability) and the equipment can be easily maintained,

- **Maintenance Prevention (MP)** starts at the design stage of a new equipment with the strategic aim at ensuring reliable equipment, easy to care for and user friendly, so that operators can easily retool, adjust, and otherwise run it,
- **Reliability entered Maintenance (RCM)** is a structured and logical process for developing or optimizing the maintenance requirements of a physical resource in its operating context to realize its "inherent reliability", where "inherent reliability",
- **Productive Maintenance (PrM)**, i.e., the most economic maintenance that raises equipment productivity,
- **Computerized Maintenance Management systems (CMMS)** assist in managing a wide range of information on maintenance workforce, spare-parts inventories, repair schedules and equipment histories,
- **Total Productive Maintenance (TPM)** is the subject of this section.

The Total Productive Maintenance (TPM) is a continuous improvement program that concerns the effective and efficient use of machines and plants. With this new approach, the responsibility for the maintenance of the systems is extended to multiple levels and it is for not only the maintenance but also, and above all, to direct operators. They are involved in the maintenance, improvement projects and simple repairs, all of which become part of their routine. For example, operators dealing daily to lubricate, clean and check the machines they use.

The productive maintenance is the most economic maintenance that raises equipment productivity. The aim of this concept is to increase the productivity of an enterprise by reducing the total cost of the equipment over the entire life, i.e., during design, fabrication, operation and maintenance, and minimizing the losses caused by equipment degradation. Productive maintenance can be consider as performing preventive maintenance, corrective maintenance, and maintenance prevention trough the entire life cycle of the equipment used in the factory. [26] [27]

The TPM is a methodology originating from Japan to support its lean manufacturing system, since dependable and effective equipment are essential prerequisite for implementing lean manufacturing. [28] It has been widely recognized as a strategic weapon for improving manufacturing performance by enhancing the effectiveness of production facilities: a production-driven improvement methodology that is designed to optimize equipment reliability and ensure efficient management of plant assets. TPM comprises:

- maximizing equipment effectiveness through optimization of equipment availability, performance, efficiency and product quality,
- establishing a preventive maintenance strategy for the entire life cycle of equipment,
- covering all departments such as planning, user and maintenance departments,
- involving all staff members from top management to shop-floor workers,
- promoting improved maintenance through small-group autonomous activities.

The emergence of TPM is intended to bring both production and maintenance functions together by a combination of good working practices, team-working and continuous improvement. TPM is about communication, it mandates that plant operators, maintenance specialists, and production engineers collectively collaborate and understand each other's language. Moreover, strategic TPM implementation can also facilitate achieving the following organizational manufacturing priorities and goals:

- **productivity**, i.e., reducing unplanned stoppages and breakdown improving equipment availability and productivity, and providing customization with additional capacity, quick change-over and design of product,
- **quality**, i.e., reducing quality problems from unstable production, reducing in field failures through improving quality, and providing customization with additional capacity, quick change-over and design of product,
- **cost**, i.e., life cycle costing, efficient maintenance procedures, supporting volume, mix flexibility, and reducing quality and stoppage-related wastes,
- **delivery**, i.e., supporting of JIT efforts with dependable equipment, improving efficiency of delivery, speed, and reliability, and improving line availability of skilled workers,

- **safety**, i.e., improving workplace environment, realizing zero accidents at workplace, and eliminating hazardous situations,
- **morale**, i.e., significant improvement in kaizen and suggestions, increasing employees' knowledge of the process and product, improving problem-solving ability, increasing in worker skills and knowledge, and employee involvement and empowerment.

In addition, TPM implementation in an organization can also lead to realization of intangible benefits in the form of improved image of the organization, leading to the possibility of increased orders. After introduction of autonomous maintenance activity, operators take care of machines by themselves without being ordered to. With the achievement of zero breakdowns, zero accidents and zero defects, operators get new confidence in their own abilities and the organizations also realize the importance of employee contributions towards the realization of manufacturing performance. [29]

The basic practices of TPM are often called the pillars or elements of TPM, and the entire edifice of TPM is built and stands on eight pillars. [30] TPM paves way for excellent planning, organizing, monitoring and controlling practices through its unique eight-pillar methodology. TPM initiatives, as suggested and promoted by Japan Institute of Plant Maintenance (JIPM), involve an eight pillar implementation plan that results in substantial increase in labor productivity through controlled maintenance, reduction in maintenance costs, and reduced production stoppages and downtimes. The core TPM initiatives classified into eight TPM pillars or activities for accomplishing the manufacturing performance improvements include (i) autonomous maintenance, (ii) focused maintenance, (iii) planned maintenance, (iv) quality maintenance (v) education and training, (vi) office TPM, (vii) development management, and (viii) safety, health and environment. [31] [32] [33] Table 2.3 depicts detailed maintenance and organizational improvement initiatives and activities associated with the respective TPM pillars.

The main tool of a TPM system is the Single-Minute Exchange of Die (SMED), i.e., a part of a larger strategy and suite of loss reduction tools that are best aligned with TPM priorities as understood through OEE losses. This resource refers to the theory and techniques for performing set-up operations, since, to meet customer demand through batch size reduction, it is important that the operator's work to ensure changeover times are reduced and changeover quality improved. Starting a

	Fostering operator skills
Autonomous maintenance	(i) Fostering operator ownership and (ii) perform cleaning, lubricating, tightening, adjustment, inspection, and readjustment on production equipment
Focused improvement	(i) Systematic identification and elimination of 16 losses, (ii) working out loss structure and loss mitigation through structured why-why and Failure Mode and Effects Analysis (FMEA), (iii) achieve improved system efficiency, and (iv) improved Overall Equipment Effectiveness (OEE) on production systems
Planned maintenance	(i) Planning efficient and effective Preventive Maintenance (PM), Predictive Maintenance (PdM), and Time-Based Maintenance (TBM) systems over equipment life cycle, (ii) establishing PM check sheets, and (iii) improving Mean-Time-To-Failure (MTBM) and Mean-Time-To-Repair (MTTR)
Quality maintenance	(i) Achieving zero defects, (ii) tracking and addressing equipment problems and root causes, and (iii) setting 3M (machine/man/material) conditions
Education and training	(i) Imparting technological, quality control, interpersonal skills, (ii) multi-skilling of employees, (iii) aligning employees to organizational goals, and periodic skill evaluation and updating
Safety, health and environment	(i) Ensure safe working environment, (ii) provide appropriate work environment, (iii) eliminate incidents of injuries and accidents, and (iv) provide standard operating procedures
Office TPM	(i) Improve synergy between various business functions, (ii) remove procedural hassles, (iii) focus on addressing cost-related issues, and (iv) apply 5S in office and working areas
Development management	(i) Minimal problems and running in time on new equipment, (ii) utilize learning from existing systems to new systems, and (iii) maintenance improvement initiatives

Table 2.3 Issues addressed by various TPM pillar initiatives.

SMED system from the perspective of a TPM culture is about operator involvement in the reduction of waste in the value stream. Becoming more agile as a result, for example being able to quickly respond to changes in the market and customer orders at short notice.

Shigeo Shingo recognizes eight techniques [34] that should be considered in implementing SMED:

1. separate internal from external setup operations,
2. convert internal to external setup,
3. standardize function, not shape,
4. use functional clamps or eliminate fasteners altogether,
5. use intermediate jigs,
6. adopt parallel operations (see image below),
7. eliminate adjustments,
8. mechanization.

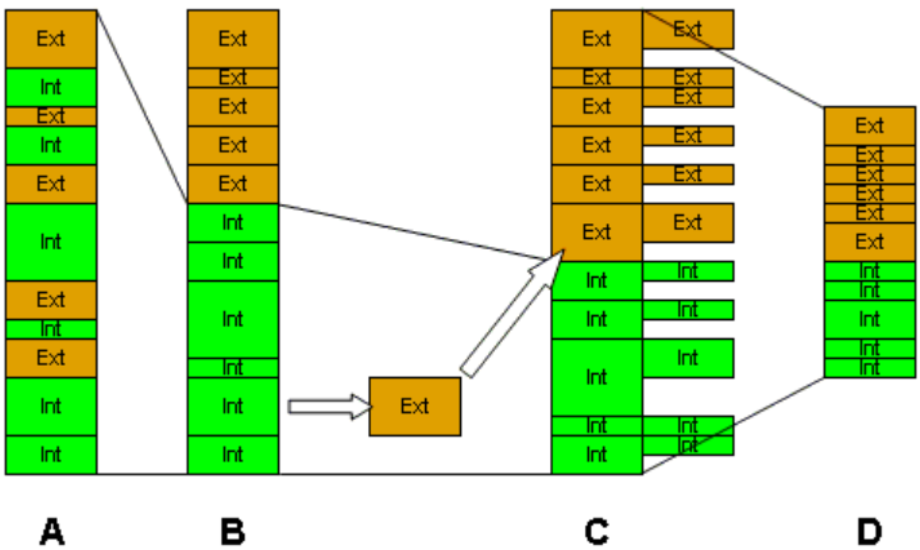

Fig. 2.2 Four implementing stages of the SMED method.

These techniques considers that an external setup can be done without the line being stopped whereas internal setup requires that the line be stopped. The same author suggests [35] that SMED improvement should pass through four conceptual stages as shown by the Figure 2.2: A) ensure that external setup actions are performed while the machine is still running, B) separate external and internal setup actions, ensure that the parts all function and implement efficient ways of transporting the die and other parts, C) convert internal setup actions to external, and D) improve all setup actions.

Getting started with SMED is straightforward, usually improvements are low cost no cost and involve operators working together with colleagues to develop standardized activities through a formal project plan, and this ensures the greatest level of success and repeatability. Contrary to the perception the best solutions are not machine redesigns and implementing high tech solutions. Work flow and workplace management coupled with visual management are the first places to start in any SMED system, there will be a lot of low hanging fruit that can provide you with quick wins.

2.1.6 Standardization, visual management, and problem solving

The value stream mapping is the graphical representation of the steps of the flow of materials and information that leads to a product from order to delivery. This tool allows you to find an immediate visual and waste, and improvement opportunities. The method involves tracking before the map of the current state (as is), and then propose changes to be included in the map of the future state (to be).

Handling and production of materials or components between successive stages of processing is authorized by the kanban (card). It provides information visual, simple and comprehensive that indicate: what, when and how much to produce, the destination of the output required, necessary materials and semi-finished products, and any other useful information. The kanban allows the continuous flow of supply, thus constituting the main instrument to implement the Just-In-Time.

2.2 Enterprise Information Systems

An Enterprise Information System (EIS) is any kind of information system which improves the functions of enterprise business processes by integration. This means typically offering high quality of service, dealing with large volumes of data and capable of supporting some large and possibly complex organization or enterprise. An EIS must be able to be used by all parts and all levels of an enterprise. [36].

EISs provide a technology platform that enables organizations to integrate and coordinate their business processes on a robust foundation. An EIS is currently used in conjunction with customer relationship management and supply chain management to automate business processes. An EIS provides a single system that is central to the organization that ensures information can be shared across all functional levels and management hierarchies. An EIS can be used to increase business productivity and reduce service cycles, product development cycles and marketing life cycles. It may be used to amalgamate existing applications. Other outcomes include higher operational efficiency and cost savings.

2.2.1 Enterprise Resource Planning (ERP) systems

Enterprise resource planning (ERP) is the integrated management of main business processes, often in real time and mediated by software and technology. ERP is usually referred to as a category of business management software—typically a suite of integrated applications that an organization can use to collect, store, manage, and interpret data from many business activities. ERP provides an integrated and continuously updated view of core business processes using common databases maintained by a database management system. ERP systems track business resources—cash, raw materials, production capacity—and the status of business commitments: orders, purchase orders, and payroll.

The Gartner Group first used the acronym ERP in the 1990s [37] to include the capabilities of material requirements planning (MRP), and the later manufacturing resource planning (MRP II), as well as computer-integrated manufacturing. Without replacing these terms, ERP came to represent a larger whole that reflected the evolution of application integration beyond manufacturing. [38] Not all ERP packages are developed from a manufacturing core: ERP vendors variously began assembling their packages with finance-and-accounting, maintenance, and human-resource components. By the mid-1990s ERP systems addressed all core enterprise functions. Governments and non–profit organizations also began to use ERP systems. [39]

2.2.2 Manufacturing Execution Systems (MESs)

The first organization which defined the tasks to be dealt by a MES was the Manufacturing Enterprise Solutions Association (MESA), a US “global community of manufacturers, producers, industry leaders, and solution providers who are focused on driving business results from manufacturing information”. MESA provided the following list of 11 functionalities [40], combined with each other, they can form a MES solution.

1. **Resource allocation and status.** Manage and monitor resources, including staff, machines, tools and make available the documents necessary to start the working operations. Further, set up the equipment, and reserve resources and dispatch orders in order to meet the target objectives.

2. **Operations/detail scheduling.** Identify the optimal sequence planning based on priorities and resources availability, in order to minimize setups and downtime.

3. **Dispatching Production Units.** Manage the flow of production units (e.g. jobs, batches or lots), and adjust it in real-time as events (e.g. reworking operations) occur on the shop-floor.

4. **Document control.** Manage and control the information significant for the production process (work instructions, drawings, specifications, environmental compliance requirements, safety instructions, etc.) as well as the "as planned" and the "as is" information. Historical data are saved; the information must be accessible to the staff at the right time and right place.

5. **Data collection/acquisition.** Data related to the production can be collected both automatically or manually, and used to track deviations.

6. **Labor management.** Provide the updated status of the personnel, store the staff working hours, the criteria to manage absences, holidays, etc, as well as

the ability to perform tasks. This package can be used to evaluate the cost of activities, and may interact with the ERP to optimize resources allocation.

7. **Quality management.** Measure production data and analyze them in real-time, aiming at ensuring product quality and identify in advance issues and criticalities. Actions to correct the issue can be included, as well as tools for process control (such as Statistical Process Control or Statistical Quality Control) and for the management of inspections and offline analyses.

8. **Process management.** Monitor the production process; alarm management functions can be included and automatic corrections or decision support tools can be integrated to correct and improve process activities.

9. **Maintenance management.** Track the use of operating material to plan periodic and preventive maintenance tasks, ensuring their availability according to the scheduled activities. The system also stores the chronology of past interventions to support problem diagnosis and the execution of maintenance actions.

10. **Product tracking and genealogy.** Record all the production data across the entire manufacturing chain, to ensure that the position of each item can be identified in real-time as well as its manufacturing history (e.g. components suppliers, lot and serial number, operators working on it, and alarms).

11. **Performance analysis.** Produce user-friendly, complete reports containing process and product information (e.g. resources availability and utilization, cycle times, and no-compliance) and a comparison with the past history and the expected performance, to support the assessment of production efficiency and the detection of issues.

Later, in the 2000s, the standard ISA 95 has been issued by the International Society of Automation. In this standard, a functional hierarchy model consisting in five levels is defined:

- level 4: business planning and logistics,
- level 3: manufacturing operations and control,
- levels 2, 1, 0: batch, continuous, discrete control.

The level 0 indicates the manufacturing process, the level 1 indicates manual sensing, sensors and actuators used to monitor the process, while the level 2 indicates the control activities that keep the process stable or under control. These tasks are not addressed in the ISA 95 standard: this document is mainly concerned with the activities for levels 3 and 4, and with their interface for data exchange. Activities in level 4 include tasks for business management, which are usually performed by an ERP. Conversely, level 3 is related to production management: the functionalities of this level correspond to the list of 11 tasks shown above. The standard ISA 95 has been adopted and extended by the International Electrotechnical Commission, named "IEC62264: Enterprise-control system integration." Currently, this standard consists in 5 parts:

1. "**models and terminology**" describes the level 3 activities, and the interfaces within level 3 and between levels 3 and 4,
2. "**objects and attributes for enterprise-control system integration**" specifies generic interface content exchanged between manufacturing control functions and other enterprise functions,
3. "**activity models of manufacturing operations management**" defines activity models of manufacturing operations management that enable enterprise system to control system integration,
4. "**objects models attributes for manufacturing operations management integration**" defines object models and attributes exchanged between the manufacturing operations management taking place in the level and previous defined,
5. "**business to manufacturing transactions**" defines business to manufacturing transactions and manufacturing to business transactions.

A schematic of the hierarchy model levels, the associated company levels, and the corresponding supporting IT tool is provided in Figure 2.3. Further, a graphical indication about the time-scales and the detail of the transmitted information is given. The time scale for the planning provided by the ERP is in the order of weeks-months; the more detailed schedule elaborated by the MES involves events in the order of hours-days; the phenomena occurring at the shop-floor have lower time-scales, in the order of minutes or hours. Conversely, on the shop-floor a huge quantity of data can be acquired. Such data must be analyzed and transformed into a smaller amount of information to be transmitted to the business level, in order to have a complete and exhaustive picture and take proper tactical decisions.

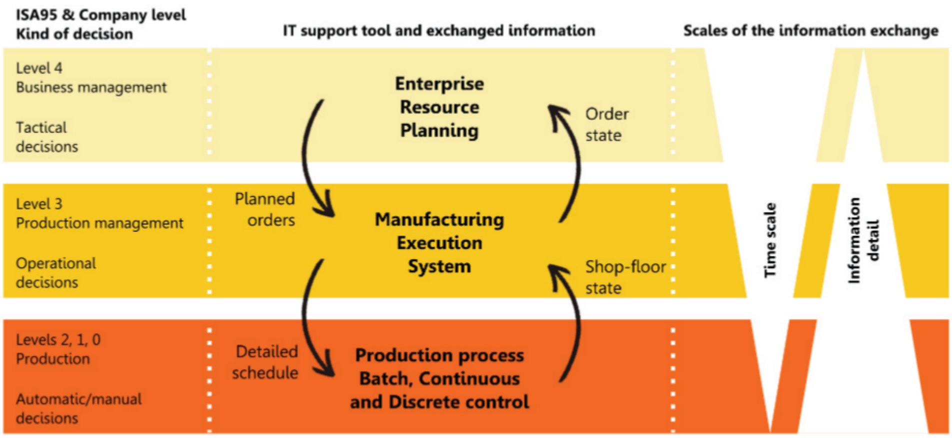

Fig. 2.3 MES positioning within an industrial framework. [41]

2.2.3 Product Lifecycle Management (PLM) systems

The Product Lifecycle Management (PLM) provides a philosophy approach and a management methodology for product development and information exchange mainly with the production and process control areas. Three main works have been analyzed to produce this section. [42] [43] [44]

It is a concept and set of systematic methods that attempts to control the product data, that are divided in:

- **definition** that are physical and functional properties,
- **life cycle** that is the sequence of the product stages in the order-delivery process,
- **metadata** that are "information about information," i.e. the description of the product data structure.

The main concepts involved in the PLM are terms and abbreviations, models of product and information, type of product or type of information object, practices and principles, related processes (product information management), and instructions.

The PLM involves a **hierarchy structure of item**, where the item is a product, an element, a module, a component, a material, a service, a document, or a software. It is an EIS that integrates the functions of the whole company as shown by Figure 2.4. According to Kenneth McIntosh, the "PLM can be the operational frame of CIM".

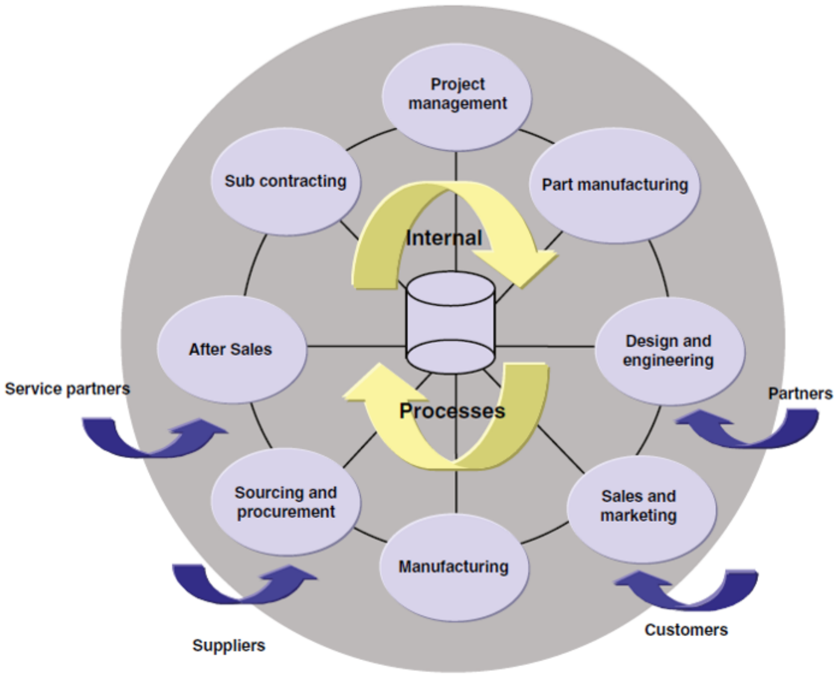

Fig. 2.4 The PLM system often creates a wide totality of functions and properties with which to support the different processes involved in the creation, recording, updating, distribution, utilization, and retrieval of information.

The PLM main features are as follow:

1. **item management**, i.e., information about the status of the item and of the processes related to the creation and the maintenance of the item,
2. **product structure management and maintenance**, i.e., managing the relationships (network structure) between items.
3. **user privilege management**, i.e., who can do what?
4. **product data storing management**, i.e., maintaining information about states, versions and changes,
5. **information retrieval**, i.e., how simplify the access to it?

6. **change management**,
7. **configuration management**, i.e., product customization,
8. **tasks management and messages**, i.e., workflow management to improve the communication,
9. **file and document management**,
10. **information loss avoiding**,
11. **backup management**,
12. **history/system log**, i.e., product process traceability,
13. **file vault**, i.e., network system and hardware.

Considering the history of the term, PLM can be defined as the new integrated business approach that, with the help of IT, implements an integrated, cooperative, and collaborative product information management along the different phases of its life cycle. In this sense, the PLM includes:

- a strategic orientation to the creation of value "on" and "through" the "product",
- the application of a collaborative approach for the valorization of the core-competences of different actors,
- the use of a consistent number of IT solutions for the practical implementation of the consequent coordinated, integrated and secure management of all the information necessary for the creation of value.

The **metadata base** is needed to maintain the structure of the whole system. The task of the metadata base is to handle relationships between individual pieces of product data, the structure of the information, and the rules and principles needed to ensure the systematic recording of the information. The **application** carries out the PLM functions of information and metadata base management and appears to the user as a variety of different user interfaces. The PLM application usually also acts as a link between different applications and systems within the sphere of PLM.

The PLM concept promises to provide support for the product's entire life-cycle, from the first conceptualization to the disposal of its last instance. The

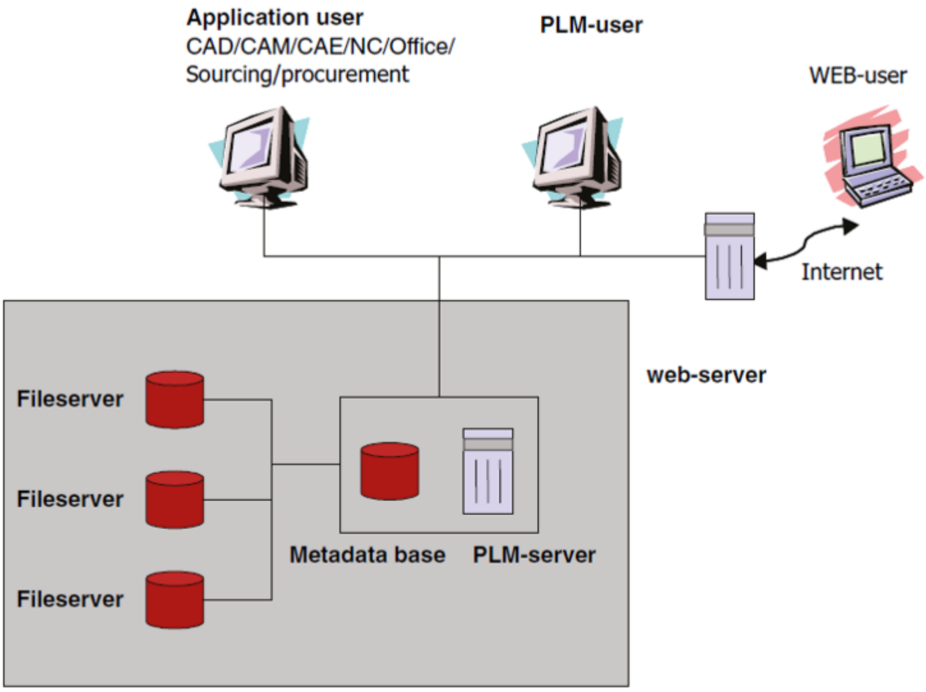

Fig. 2.5 Example of a PLM system architecture.

volume, diversity, and complexity of information describing the product will increase correspondingly. In the literature, there are many proposals of framework for product information management that can access, store, serve, and reuse all the product information throughout the entire life-cycle. Research is needed to identify and model all the necessary components of an effective framework of PLM that includes standards and conceptual Application Program Interface (API) between all such components and the other information systems involved in the product life-cycle (and, generally speaking, in the manufacturing CPS of a plant). It is clear that design and manufacturing process are the main functionalities for a sustainable production: such components inter-operate by exchanging product information clearly associated with a Big Data environment, and such product information modeling frameworks need to support such "horizontal" information exchanges as readily as the "vertical" exchanges among other PLM systems, components of a MES, and any intermediary systems, such as Product Data Management and ERP systems. [45]

In a concise definition, PLM is the commercial activity of managing the products of a company throughout its life-cycle, from the first idea of a product until it is removed and discarded most effectively. [43] Further, PLM may mean a technological solution that encompasses different and complementary tools to promote collaboration between stakeholders in order to support the product life-cycle's effective management. Although PLM concepts are more commonly found in traditional manufacturing companies, their concepts can also be found in other business profiles such as startups. [46]

Concluding, considering the 4.0 era, it seems that the overlap, or a more comprehensive integration of the Fourth Industrial Revolution, Sustainable Development to Product Lifecycle Management, may promote innovation, more robust outcomes in a long-term business perspective. Research efforts are dedicated to find elements that can support the use of 4.0 technologies in Product Lifecycle Management with a strong background in Sustainable Development, concept that will soon be integrated into a 5.0 perspective. Examples of forces driving the 4th Industrial Revolution to PLM with sustainability purposes identified by the literature are concepts and approaches like flexibility (which means focusing without losing flexibility), durability and Big-Databases that may increase the product lifetime, reduce the amount of the product in circulation in combination with maintenance and refurbishment operations. [47] Further, 4.0 technologies allow the identification of new markets and the product will be more competitive in the market and more profitability for the enterprises who adopt this type of technology approach. The sustainability is a crucial factor for innovation, competitiveness and survival in these uncertain times. However, there are many barriers to integrating technologies, which have their implementation challenges when considered isolated. Lack of information and knowledge about technologies and sustainability, availability of appropriated equipment, human resources engagement and lack of skills are some of the main challenges faced by PLM for Sustainability adopters. [48] A systematic review in the literature selects 12 works in the literature texts that are successful examples where every enterprise has been ready to choose the work method that better fits its profile. [47]

2.2.4 Discussion about ERP-MES-PLM integration

One of the pillars of Industry 4.0 is the integration among data coming from different systems. The machines of a manufacturing plant are starting to be connected, following the Industrial Internet of Things (IIoT) paradigm. Currently, the usage of IoT is limited to analyze shop floor behavior by monitoring the environment parameters. [49] However, the most significant challenge is integrating the main IT management systems towards a Horizontal and Vertical integration expressed by the concept of Industry 4.0. Particularly, the integration of PLM, ERP and production systems guarantee the full control over the whole product life-cycle and among all functional offices of the company. It's an end-to-end process where data will flow seamlessly across different departments within the company, thus breaking the walls between the different functional areas in a company. An example of an architecture of integration between PLM, MES and ERP is given by a central system called Knowledge Base System (KBS). KBS consists, in this example, of a database in addition to the individual databases of each system. It is a structured central database in which the IT systems can transfer and withdraw the necessary information. It's a bi-directional data flow between the IT systems and the KBS as shown in Figure 2.6. The proposed architecture will allow to collect all the information related to the new component in the KBS system, which helps to perform analysis on a product data. This analysis of product data helps in finding the patterns of different products of a same family, which will reduce the time for the product and process design and able to monitor the performance in the production line. To design such architecture, we have to observe the functional framework in one-of-a-kind production and the data flows among the PLM, MES and ERP.

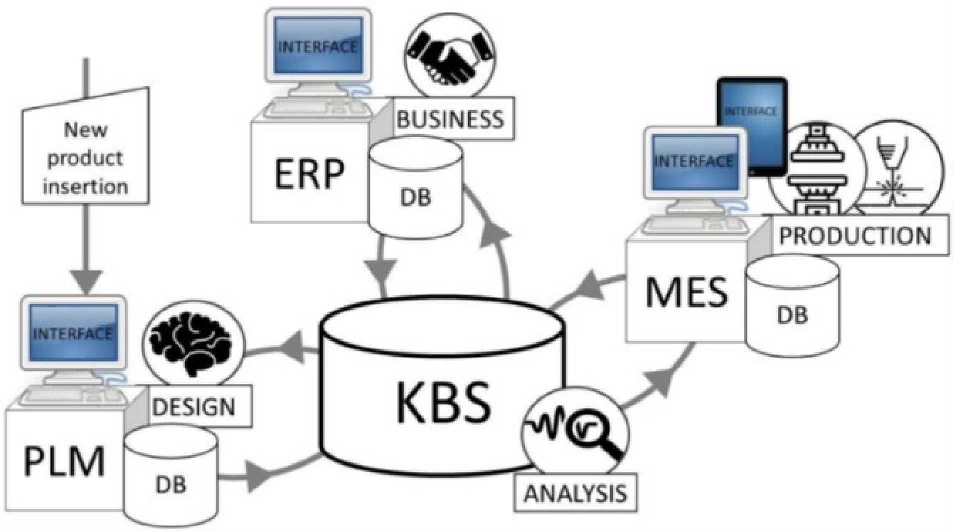

Fig. 2.6 Conceptual architecture of the knowledge base system. [11]

The generic functional framework in OKP is shown in Figure 2.6. The Customer Relationships Management (CRM) module of the ERP systems manages pre-sales activities and sends information about to the PLM system in order to accept, refuse of plan the shop order comparing similar product history. This information is sent by the ERP to and retrieved by the PLM from the KBS in order to reducing the time in product planning phase and improving the cost estimation in OKP. When the product is defined, and when a production order is approved, the PLM system provides the best receipt (processes description) for the product and share the information with the ERP system and the MES. The MES gives the order command to the shop floor to build a prototype and checks for any design modifications and sharing such information with the PLM trough the KBS until the prototype meets all the quality requirements. At this moment the MES provides the bill of materials (BOM) information to the ERP for the resources allocation powered by the Material Requirement Planning (MRP) and Supplier Resources Management (SRM) modules in order to satisfy the customer demand. Once everything is setup, production process will start and the final product will be controlled by ERP system for the delivery to the customer. KBS as a system gives continuous feedback to the MES for any performance improvement in production line. [11]

The KBS, basically, is a database with APIs that makes possible the communication among the systems and provides a PLM-MES (product-process) information that represents an extract mainly of the one stored in MESs and PLM ones like shown in Figure 2.7. ERP system generates product specifications and stores data in the entity Product. Product Model(s) entity, which is a PLM item, generates one or more models for this product and each model consists of one or more parts (BOM). Manufacturing Process Plan is a PLM entity that defines the parameters of a production cycle for a product model and the corresponding entity called Manufacturing Operations defines the kind of operation, required resources and machines. It shares the relationship with list of operations entity which contains predefined operation details. On the other hand, MES entities called Production Request and Production Planning, help to define the production cycle and can be divided into one or more secondary orders. Production Request contains the data of a production orders and Production Planning consists of data to manage the programming of operations of an individual cycle relative to a particular order. One of the key entity of the MES is Production Status because it contains information about the progress of the production. This entity evaluates the data en-tries to the Physical machine entity. Each one of the Physical Machine entity can have associated at most one Production Status. The Check Start Output entity is linked to the Check Start Machine, which imposes certain controls to monitor before using a machine. Final product entity provides information about the status of a set of by-products and/or final products. This entity evaluates the statements saved in Production Status at time intervals, obtaining as a result the status of the articles declared: good, scrap or re-cycle. Final delivery entity, which is a ERP item, stores the data of final products labeled as good and delivered to the customer.

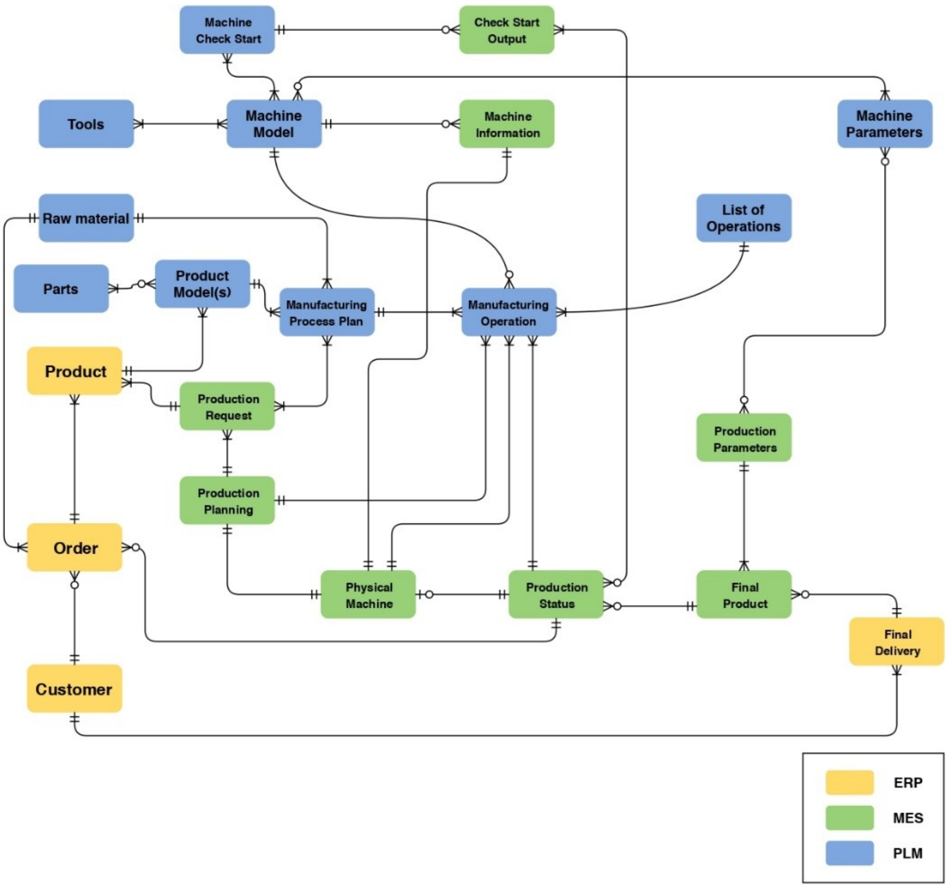

Fig. 2.7 Entity-Relationship diagram in KBS DB. [11]

Specific APIs are developed in order to allow the communication between such systems: the connectivity is crucial to implement the data flow, the Internet of Things is all about connectivity and API is a central concept of this technology. Currently REST and MQTT API's are predominantly in use. REST (REpresentational State Transfer) is designed as request/response that communicates over HTTP and MQTT (Message Queuing Telemetry Transport) is designed as publish/subscribe that communicates over TCP/IP sockets or WebSockets. In REST, requests made to an application URL and will obtain a response with a payload formatted in HTML, XML, JSON, or some other format. The REST API is a request/response model, i.e., it is possible to set the repeat mode to an entity based-on time frame for the request/response depending on the priority. For example, Machine model or Machine Check start entities in PLM are static data to store in KBS which requires less priority compare to Manufacturing Process Plan or Manufacturing Operations. Other works on the integration of such systems were developed by the author of this paper and the research team in order to investigate technological and management limitations and opportunities. [10] [11] [12] [13] [14]

2.3 Industry 4.0 and Cyper-Physical Systems (CPSs)

In recent years, Industry 4.0 has attracted great attention from both manufacturing companies and service systems. On the other hand, there is no certain definition of Industry 4.0 and naturally, there is no definite utilization of the emerging technologies to initiate the transformation of Industry 4.0. Mainly, Industry 4.0 is comprised of the integration of production facilities, supply chains and service systems to enable the establishment of value added networks. Thus, emerging technologies such as big data analytics, autonomous (adaptive) robots or collaborative robots (cobots), cyber physical infrastructure, simulation, horizontal and vertical integration, Industrial Internet, cloud systems, additive manufacturing and augmented reality are necessary for a successful adaptation. The most important point is the widespread usage of Industrial Internet and alternative connections that ensure the networking of dispersed devices. As a consequence of the developments in Industrial Internet, in other words Industrial Internet of Things (IIoT), distributed systems such as wireless sensor networks, cloud systems, embedded systems, autonomous robots and additive manufacturing have been connected to each other. Additionally, adaptive robots and CPS provide an integrated, computer-based environment that should be supported by simulation and three-dimensional (3D) visualization and printing. Above all, entire system must involve data analytics and miscellaneous coordination tools to conduct a real time decision making and autonomy for manufacturing and service processes. While constructing the framework, network of sensors, real-time processing tools, role-based and autonomous devices are interpenetrated with each other for real-time collection of manufacturing and service system data. For successful implementation of Industry 4.0 transformation, three core and nine fundamental technologies are required to be the part of the entire system.

2.3.1 Additive Manufacturing, mass customization and 4.0 design

Additive Manufacturing (AM) is a set of emerging technologies that produces three dimensional objects directly from digital models through an additive process, particularly by storing and joining the products with proper polymers, ceramics, or metals. In details, additive manufacturing is initiated by forming Computer-Aided Design (CAD) and modeling that arranges a set of digital features of the product and submit descriptions of the items to industrial machines. The machines perform the transmitted descriptions as blueprints to form the item by adding material layers. The layers, which are measured in microns, are added by numerous of times until a three-dimensional object arises. Raw materials can be in the form of a liquid, powder, or sheet and are especially comprised of plastics, other polymers, metals, or ceramics. [50] In this respect, AM is comprised of two levels as software of obtaining 3D objects and material acquisition side.

AM is emerging as an important manufacturing process and a key technology for developing innovative products. However, in order to support and promote this complex but often beneficial technology, adequate EISs entirely dedicated to this manufacturing process are needed. During AM, a large amount of data is generated, exchanged, and used from which to draw information about materials, design patterns, processes, and measurements. This is where process knowledge comes from: in fact, while data about a single part is essential for its traceability, if collected in a methodical way, the history of thousands of components can be analyzed to streamline or automate all phases of the part lifecycle. A scenario that enhances even more the ability to manage such data is Mass Customization (MC): [51] [52] [53] a scenario in which AM plays a key role. To make this strategy sustainable, in addition to the use of more flexible technologies, it is necessary to have an integrated management of all the information produced by the company and the ecosystem in which it operates.

The latest technological developments introduced by Industry 4.0 have increased the amount of information linked to a production process or value stream, from obtaining raw materials to a saleable product. The figure 2.8 shows the traditional product design paradigm in opposition to the one that supports mass customization, where customer requirements are considered after the design process. Thanks to this different management of the information provided by the customer, it is possible to

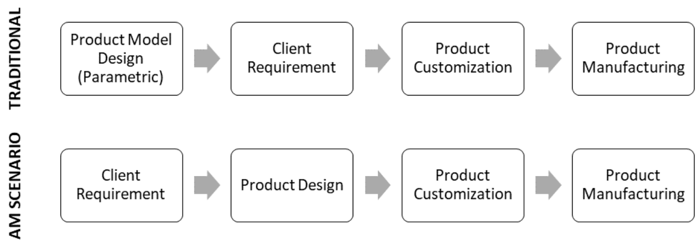

Fig. 2.8 Traditional product design paradigm.

change the product model enough to have a customized good but avoiding the design of the model from scratch. The figure 2.9 shows the mapping of the information flow during the customization process supported by additive manufacturing. The company's tasks begin with the development of the general parametric model containing both the bill of materials and the parametric CAD model, as well as the machining and optional assembly components. The customer comes into play with the supply of the specific parameters of the individual product, after which he approves or rejects the variant proposed by the company based on price and delivery time: (a) rejection leads to the creation of a further variant or suspension of the order, while, (b) with an approval, on the other hand, the process continues with the upload of the final variant data through the PLM platform and a G-code file of the model is generated to produce the part and complete the order.

A shared database, capable of supporting design, production, and all life cycle processes, is a complex resource to design, develop and maintain. In an additive manufacturing scenario, this complexity may be even greater due to the volume and low repeatability of process and quality data. In order to deal with the criticality of the management of the information baggage that characterize the AM, an information system that is based on the product life cycle, like the PLM, is therefore required. It has three basic functional objectives:

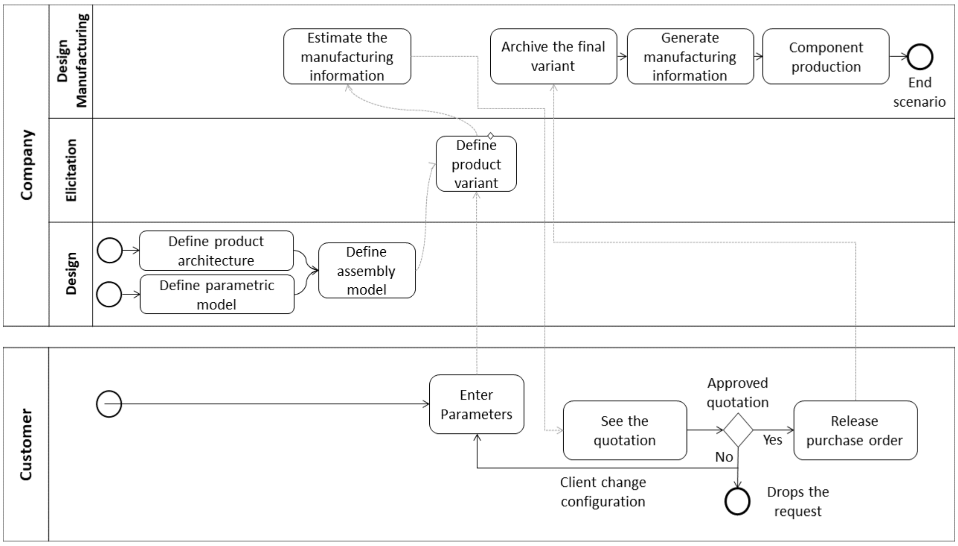

Fig. 2.9 Map of the information flow during the customization process supported by additive manufacturing.

- to offer a standardized protocol for data interoperability in order to share relevant information as each product continues in its life cycle
- integrate all life cycle processes by providing a consistent flow of data within a heterogeneous framework
- coordinate information and data in an open, lightweight, and extensible form.

For the development of an AM plan involving several 3D printers, it is necessary to use a multi-layer model due to the many interconnected processes. The figure 2.10 shows an example of an architecture in which several modules of the most common ERP solutions on the market are considered. Once the customer order is accepted, the available production resources are determined, including employees, materials and production equipment. In order to obtain information on the required production resources, a machine code is generated, based on the type of production facility, so the production process is simulated to determine how many resources are required and the duration of the process. This data is linked to the CAD model and the specific machine, as production plants of the same type can have different parameters. A revision of a product coincides with a complete new input of the production resources required for the model. Requirements for production are calculated from the bill of materials and sales orders. Production is planned by optimizing the level of machine utilization. The grouping module handles a set of components to be completed simultaneously as required for assembly. Parts can be assigned to the same machine with the help of a nesting algorithm or distributed to different printers and the start-up of a production plant is managed through the MES.

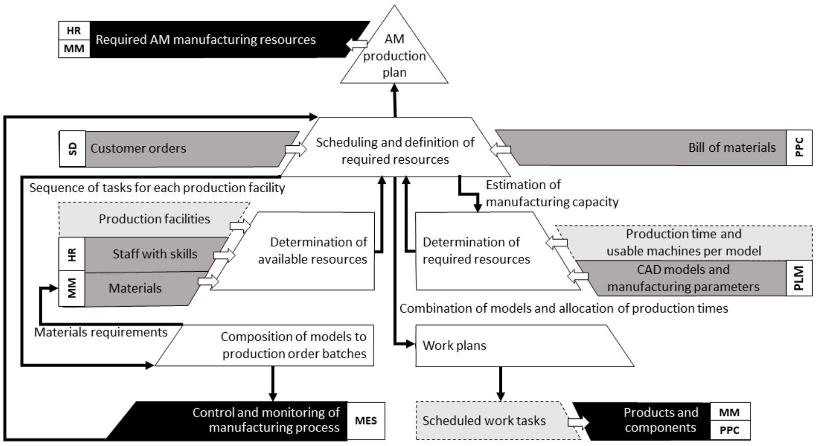

Fig. 2.10 Example of an ERP architecture solution dedicated to the development of an AM plan involving several 3D printers

The figure 2.11 shows schematically a possible integration structure of MES with the Design For Additive Manufacturing (DFAM) process. The MES supports the development of prediction and simulation models that are based on and validated by feedback from a set of sensors. The improved models thus better support DFAM, for example, in decision support for part orientation, which has a strong impact on quality. Secondly, information from the laboratory floor can also be used for further adjustments of the part geometry. Finally, MES supports the optimization of manufacturing routes and machine parameters through the use of histories of different variables that may influence the manufacturing process, concerning deposition and material. Thus, MES can improve the additive process, contribute to knowledge creation in a mass customization scenario and streamline the trial and error process. Improving the knowledge process results in better predictability and quality of output. This, in turn, can change the economic balance between traditional and additive manufacturing, making the latter cost-effective even for larger batch sizes. In addition, sensors can be used to monitor parameters related to the energy impact of the process, such as nozzle temperature and energy required. Giving a last consideration about MESs, it should be noted that the use of MES could be much more significant in high-value productions, such as the production of aeronautical, bio-mechanical or jewelry components.

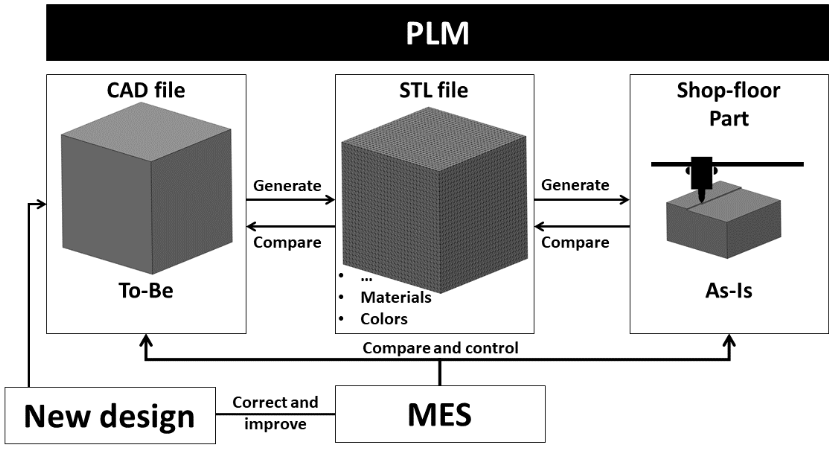

Fig. 2.11 Example of integration structure of MES with the Design For Additive Manufacturing (DFAM) process.

The development and the industrial adoption of additive technologies is leading to a scenario in which the entire product development chain is digitized: from the initial design of products to the simulation and optimization of manufacturing processes and the maintenance of assets and products using sensor systems. One of the challenges brought by the organization of the AM value chain is the development of information systems that can help decision making but also support data management in a consistent and coherent way. To support decision making and address data management issues, a formal, shared, and explicit representation of expert knowledge and processes is required, whose semantics can be processed by information systems: a widely used example is given by ontologies. [54] The Guide to Principal to Rule (GPR) is the main guide to the development of design rules in additive manufacturing on which several ontologies have been based, formalizing the use of concepts, primitive concepts, modules, and rules. Focusing on the analysis of manufacturing, it is possible to structure factory knowledge into three related categories as in the figure 2.12, which shows a possible ontology structure that maintains the three components of the GPR. Although it is a simple example, this knowledge representation results in a business asset that can:

- promote the application of design principles,

- reduce ambiguities and inconsistencies,
- support the creation of new principles,
- and model knowledge as a set of modular components in relation to each other.

In the figure 2.12, the basic entities, i.e., concepts, are represented as rectangles and grouped into macro-modules used to formulate product and process design rules. Arrows indicate the relationships between the concepts. Product design is the definition of the design features and geometric parameters of the component. Process design, on the other hand, is the process and material parameters. The term manufacturing features refers to the intermediate class between product and process design that is responsible for modeling manufacturing constraints such as tolerances, thicknesses, space constraints or support structures. They represent a description of the admissibility domain of additive processes as the geometric regions of the component change. KBSs in the additive industry store information in databases designed often using ontologies like the one discussed. It is critically important to design a database with thinking about how to manage factory knowledge from which knowledge can be derived to support business decisions, whether dictated by human insights or artificial intelligence algorithms.

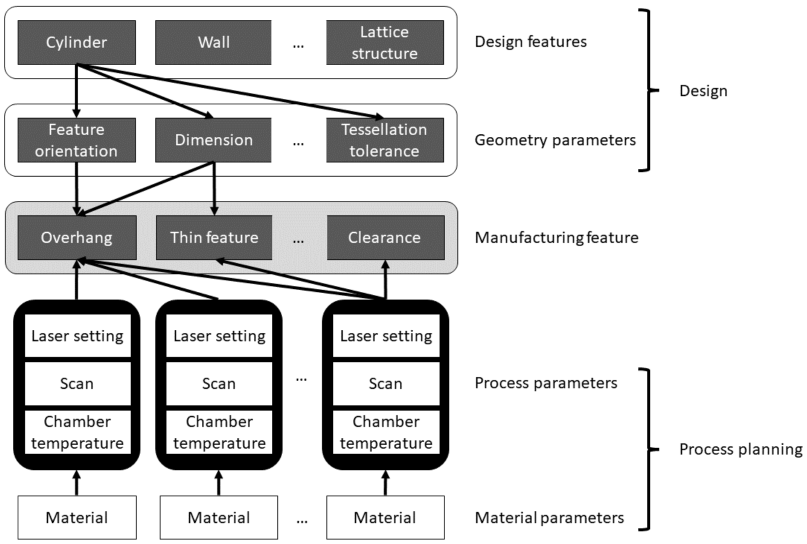

Fig. 2.12 .

Today, companies are increasingly trying to improve the organization and capitalization of their informative assets and, in particular, the know-how of their human resources. This knowledge management activity, in fact, has become a crucial component for the sustainability of a company in today's Industry 4.0 context. The additive industry, to reap the benefits of its processes, is forced to deal with increasingly complex product design, process optimization and production management. For example, in product design, Mass Customization, a favorable scenario for the AM, requires a deep knowledge of the dependencies between the variables involved for a rapid response to continuous and increasingly complex changes in product requirements. Assuming sufficient information to obtain the required knowledge, the knowledge extraction process is done through specifically designed information system components. The main requirement for such components is to simplify user access to information in many cases until generating knowledge not directly required. A component designed with this purpose potentially supports the wisdom of an enterprise, i.e., supports the decisions of the human resources of that enterprise or, especially for operational decisions such as nesting activities, makes those decisions automatically by replacing the human being.

Machine Learning (ML), therefore, can support human decisions and provides a knowledge base on which to set up automated decisions based primarily on concepts specific to operations research. Overall, ML has had a positive impact on the prospects for adopting AM and improving its value proposition. That said, most of the ML applications for AM are not yet robust or reliable enough to be adopted in industry but the advancement in research is rapid and bodes well for rapid deployment of these technologies as additive process adoption increases. Artificial Intelligence (AI) models, as well as being an essential tool to support human decisions, can be developed in order to automate business decisions, which consists in automating the entire process that goes from the acquisition of data until the transformation of the extracted information from them into a physical action impacting the process.

OKP

design montagna

2.3.2 Main technologies

Adaptive robotics (cobots) are robots intended for direct human robot interaction within a shared space, or where humans and robots are in close proximity. As a consequence of the combination of microprocessors and AI methodologies, the products, machines and services become smarter in terms of having not only the abilities of computing, communication, and control, but also having autonomy and sociality. In this regard, adaptive and flexible robots combined with the usage of artificial intelligence provide easier manufacturing of different products by recognizing the lower segments of each parts. This segmentation proposes to provide decreasing production costs, reducing production time and waiting time in operations. Additionally, adaptive robots are useful in manufacturing systems especially in design, manufacturing and assembly phases. [55] For instance, assigned tasks are divided into simpler sub-problems and then are constituted a set of modules in order to solve each sub-problem. At the end of each sub task completion, integration of the modules to reach an optimal solution is essential. One of the sub technologies underlying adaptive robots can be given from co-evolutionary robots that are energetically autonomous and have scenario based thinking and reaction focused working principle. [56]

Cyber-Physical System (CPS) can be explained as supportive technology for the organization and coordination of networking systems between its physical infrastructure and computational capabilities. In this respect, physical and digital tools should be integrated and connected with other devices in order to achieve decentralized actions. In other words, embedded systems generally integrate physical reality with respect to innovative functionalities including computing and communication infrastructure. [57] In general, an embedded system obtains two main functional requirements: (i) the advanced level of networking to provide both real-time data processing from the physical infrastructure and information feedback from the digital structure, and (ii) the intelligent data processing, decision-making and computational capability that support the physical infrastructure. [58] For this purpose, embedded systems consist of RTLS technologies, sensors, actuators, controllers and networked system that data or information is being transformed and transferred from every device. In addition to that, information acquisition can be derived from data processing and data acquisition in terms of applying computational intelligence supported by learning strategies such as case based reasoning.

Cloud based operating is another essential topic for the contribution of networked system integration in Industry 4.0 transformation. The term “cloud” includes both cloud computing and cloud based manufacturing and design. Cloud manufacturing implies the coordinated and linked production that stands “available on-demand” manufacturing. Demand based manufacturing uses the collection of distributed manufacturing resources to create and operate reconfigurable cyber-physical manufacturing processes. Here, main purpose is enhancing efficiency by reducing product life-cycle costs, and enabling the optimal resource utilization by coping with variable-demand customer focused works. [59] [60] Comprehensively, cloud based design and manufacturing operations indicate integrated and collective product development models based on open innovation via social networking and crowd-sourcing platforms. [59] [60] As a consequence of the advancements in cloud technologies such as decreasing amount of reaction times, manufacturing data will increasingly be practiced in the cloud systems that provide more data-driven decision making for both service and production systems. [61]

Virtual Reality (VR) and **Augmented Reality (AR)** are called virtualization technologies and they are entitled the integration of computer-supported reflection of a real-world environment with additional and valuable information. [62] In other words, virtual information can be encompassed to real world presentation with the aim of enriching human's perception of reality with augmented objects and elements. [63] For this purpose, existing VR and AR applications associate graphical interfaces with user's view of current environment. The essential role of graphical user interfaces is that users can directly affect visual representations of elements by using commands on appeared on the screen and interacts with these menus referenced by ad-hoc feedbacks.

Simulation is the imitation of the operation of a real-world process or system over time. Before the application of a new paradigm, system should be tested and reflections should be carefully considered. Thus, diversified types of simulation including discrete event and 3D motion simulation can be performed in various cases to improve the product or process planning. [64] For example, simulation can be adapted in product development, test and optimization, production process development and optimization and facility design and improvement.

Data analytics and **Artificial Intelligence (AI)** express the use of techniques and methodologies for massive utilization of the large masses of available data. In consequence of the manufacturing companies start to adopt advanced information and knowledge technologies to facilitate their information flow, a huge amount of real-time data related to manufacturing is accumulated from multiple sources. The collected data which is occurred during R&D, production, operations and maintenance processes is increasing at exponential speed. [65] In particular, data integration and processing in Industry 4.0 is applied for improving an easy and highly scalable adaptation for data flow based performance analysis of networked machines and processes. [66] Data appears in large volume, needed to be processed quickly and requires the combination of various data sources in diversified formats. For instance, data mining techniques have to be used where data is gathered from various sensors. This information assists the evaluation of current state and configuration of different machinery, environmental and other counterpart conditions that can affect the production as seen in smart factories. The analysis of all such data may bring significant competitive advantage to the companies that they are able to be meaningfully evaluate the entire processes. [67]

Communication and networking or **Industrial IoT (IIoT)** can be described as a link between physical and distributed systems that are individually defined. Using communication tools and devices, machines can interact to achieve given targets, focus on embedding intelligent sensors in real-world environments and processes. IIoT relies on both smart objects and smart networks and also enables physical objects integration to the network in manufacturing and service processes. In other words, major aim of IIoT is to provide computers and machines to see and sense the real world applications that can provide connectivity from anytime, anywhere for anyone for anything. Considering manufacturing advancements supported by communication and networking technologies, manufacturing industries are ready to improve the production processes with big data analytics to take the advantage of higher compute performance with open standards and achieve the availability of industry know-how in advance. [68] As a result of the penetration of manufacturing intelligence, manufacturers can be able to enhance quality, increase manufacturing output. This knowledge provides better insights for detecting root cause of production problems and defect mapping, monitor machine performance and reduce machine failure and downtime. Therefore, IIoT or communicative systems are not only considered as a technology of Industry 4.0 but also evaluated as a "cover" that contains many features from Industry 4.0 tools.

These technologies require a fundamental structure for the successful implementation. Therefore, RTLS and RFID technologies, cyber security, sensors and actuators and mobile technologies are the infrastructure for supportive technologies.

RTLS and RFID technologies mean real-time location system and radio frequency identification. Smart Factory has some critical operations such as smart logistics, transportation and storage by satisfying efficient coordination of embedded systems and information logistics. These operations include identification, location detection and condition monitoring of objects and resources within the organization and across company using Auto-ID technologies. The aggregation and processing of the real time data gathered from production processes and various environmental resources assist the integration of organization functions and enables self-decision making of the machines and other smart devices. Thus, RFID and RTLS may generate value in manufacturing and logistics operations [69] as described the basic concepts of real time monitoring systems in the following way:

- identification—especially RFID with single and bulk reading,
- locating—RTLS like GPS and others,
- sensing, e.g. temperature and humidity sensors.

Cyber security is the protection of computer systems and networks from information disclosure, theft of, or damage to their hardware, software, or electronic data, as well as from the disruption or misdirection of the services they provide. As mentioned in previous sections, Industry 4.0 transformation requires intensive data gathering and processing activities. Thus, security of the data storage and transfer processes is fundamental concepts for companies. The security should be provided in both cloud technologies, machines, robots and automated systems considering the following issues:

- data exportation technologies security,
- privacy regulations and standardization of communication protocols,
- personal authorization level for information sharing,
- detection and reaction to unexpected changes and unauthorized access by standardized algorithms.

To avoid the results of these issues, operational recovery, end-user education, network security and information security should be ensured by cyber incident response, critical operation recovery and authorization level detection programs. Other preventive actions can be access controls of user account, firewalls, intrusion detection systems and penetration tests that use the vulnerability scanners.

Sensors and actuators are the basic technology for embedded systems as entire system obtains a control unit, usually one or more micro-controller, which monitor the sensors and actuators that are necessary to interact with the real world. In industrial adaptation of Industry 4.0, embedded systems similarly consist of a control unit, several sensors and actuators, which are connected to the control unit via field buses. The control unit conducts signal processing function in such systems. As smart sensors and actuators have been developed for industrial conditions, sensors handle the processing of the signal and the actuators independently check production current status, and correct it, if necessary. These sensors transmit their data to a central control unit, e.g. via field buses. In this respect, sensors and actuators can be defined as the core elements for entire embedded systems. [70]

Mobile technologies made a significant progress after these devices were first introduced and are now so much more than just basic communication tools. These devices ensure the internet enabled receiving and processing of large amounts of information and are provided with high quality cameras and microphones, which again allow them to record and transmit information.

2.3.3 Autonomy and digital twin

With the coming of big data-driven manufacturing era, many new technologies, such as internet of things (IoT), big data, service-oriented technology, and cloud computing, have been employed in PLM. However, the current technologies mainly focus on physical product data rather than the data from virtual models. On the one hand, data generated in various phases of the whole product lifecycle may form the information island between different phases of product lifecycle. And on the other hand, a lot of duplicate data exists in different phases of product lifecycle and leads to resources waste and data sharing inefficiency. Besides, the interaction and iteration between big data analysis and various activities in the whole product lifecycle are relatively absent. To solve the problems, digital twin, with the characteristics of ultra-high synchronization and fidelity, convergence between physical and virtual product, etc., has high potential application in product design, product manufacturing, and product service. [71] [72]

In order to be able to respond quickly to unexpected events without central re-planning, future manufacturing systems will need to become more autonomous.

Autonomous systems are intelligent machines that execute high-level tasks without detailed programming and without human control. They know their capabilities (that are modeled as "skills") and their state. They are able to decide between a set of alternative actions, orchestrate and execute skills. In order to make this happen, the autonomous systems will need access to very realistic models of the current state of the process and their own behavior in interaction with their environment in the real world – typically called the "Digital Twin". [16] Autonomy provides the production system with the ability to respond to unexpected events in an intelligent and efficient manner without the need for re-configuration at the supervisory level. Lastly, ubiquitous connectivity such as the Internet of Things facilitates closing of the digitalization loop, allowing next cycle of product design and production execution to be optimized for higher performance.

The concept of using "twins" is rather old. It dates back to NASA's Apollo program, where at least two identical space vehicles were built to allow mirroring the conditions of the space vehicle during the mission. One vehicle remaining on earth was called the twin. The twin was used extensively for training during flight preparation. During the flight mission it was used to simulate alternatives on the earth based model, where the available flight data were used to mirror the flight conditions as precise as possible, and thus to assist the astronauts in orbit in critical situations. In this sense, every kind of prototype, which is used to mirror the real operating conditions for simulation of the real time behavior, can be seen as a twin. [73] [74]

As a key enabling technology with the characteristics includes interactive feedback between cyberspace and physical space, data fusion and analysis, iterative optimization for decision-making, the digital twin has been a research hot-spot of intelligent manufacturing. [75] [72] [71] In the scientific literature, searching for papers with the expressions digital twin and wisdom in the title, keywords, or abstract, there are only 10 articles about this issue and all published from 2020 onward. It is clear from this work that the digital twin is a technology that involves more than just a set of knowledge regarding a physical resource. For the digital twin, therefore, it is necessary to structure an architecture that involves a layer referring to decision making as well as the knowledge necessary to be able to make those decisions with as much awareness as possible. Digital twin is an important technology for realizing concepts such as digitalization, intelligence, and service. It integrates the attributes of multi-physics, multi-scale and multi-discipline, has real-time synchronization, faithful mapping, high fidelity and other characteristics, can accurately reflect and predict the real state of the physical world, and predict future development trends in advance. [76] Organizations are intensely developing digital twins to correctly and efficiently answer questions about the history and behaviour of physical systems. However, it is not clear how to construct these infrastructures starting from the data, information, knowledge, and wisdom available in the organization. [77]

2.4 Industry 5.0

2.4.1 European Union view

The European Commission published from the 2020 three interested article to understand the European vision concerning the Industry 5.0. [78] [79] [80]

The vision is based on three main aspects: European Green Deal, Europe Fit for the Digital Age and an Economy that Works for People. It recognises the power of industry to achieve societal goals beyond jobs and growth, to become a resilient provider of prosperity, by making production that respects the boundaries of our planet and placing the well-being of the industry worker at the centre of the production process. It complements the existing "Industry 4.0" paradigm by having research and innovation drive the transition to a sustainable, human-centric and resilient European industry. It moves focus from solely shareholder value to stakeholder value, for all concerned.

Applying ever-more advanced digital technologies. Sensor technologies, big data and AI are increasingly automating, interconnecting and optimising a wide range of industrial processes. Industry 4.0 is primarily a techno-economic vision, indicating how more general technological advancements, often originated in non-industrial contexts, will be brought to bear on industrial value chains and how they will change industry's economic position. It describes how industry will use technology to cope better in a changing world and economy, and we believe it does this very well. A transformed industry will have a transformative impact on society as well. Changing roles and increased reliance on complex technologies will require new skills. Will workers be empowered in their industrial work and attracted to work in new high-tech environments? Innovation will require industry to re-think its position and role in society. A renewed European "Industry 5.0" needs to be characterized by industries more future-proof, resilient, sustainable and human-centred with emerging drivers for the industry of the future, making an emphasis on the perspective of the industry worker. We do not distinguish between "blue collar" and "white collar" workers; in Industry 5.0, the lines between different types of industry workers are blurred. Furthermore, European values and fundamental rights should be binding principles, including respect for privacy, autonomy, human dignity and general workers' rights. It emphasises aspects that will be deciding factors in placing industry in future European society; these factors are not just economic or technological in nature, but also have important environmental and social dimensions.

The need is to better integrate social and environmental European priorities into technological innovation and to shift the focus from individual technologies to a systemic approach. Six categories have been identified, each of which is considered to unfold its potential combined with others, as a part of technological frameworks:

- individualised Human-machine-interaction;
- bioinspired technologies and smart materials;
- digital twins and simulation;
- data transmission, storage, and analysis technologies;
- Artificial Intelligence;
- technologies for energy efficiency, renewable, storage and autonomy.

The Internet of Things (IoT) is going to have a significant impact on the organisation of production thanks to a new interplay between humans and machines and a new wave of digital application to manufacturing. Industry 4.0 has focused less on the original principles of social fairness and sustainability, and more on digitalization and AI-driven technologies for increasing the efficiency and flexibility of production. The concept of Industry 5.0 provides a different focus and highlights the importance of research and innovation to support industry in its long-term service to humanity within planetary boundaries. The concept of the Society 5.0 essentially takes the digitalization and transformation dimensions, mainly situated on the level of individual organisations and parts of society, to a full national transformational strategy, policy and even philosophy. The Society 5.0 attempts to balance economic development with the resolution of societal and environmental problems. It is a society in which advanced IT technologies, Internet of Things, robots, artificial intelligence and augmented reality are actively used in every day life, industry, healthcare and other spheres of activity, not primarily for economic advantage but for the benefit and convenience of each citizen. The recently published White Paper on a regulation of artificial intelligence, as well as the European Data Strategy, clearly illustrate the importance the European Commission attaches to the societal impact of digital technologies. In particular, the Radical Innovation Breakthrough Inquirer (RIBRI) report, which identified 100 potential innovation breakthroughs.

Several Horizon 2020 funded projects have developed evidence and further guidance on the transformative elements pertinent to Industry 5.0: they develop solutions that render the production more sustainable, resilient and competitive on a long-term basis, and tackle challenges associated with beneficial human-machine interaction and skills matching. A growing number of projects is addressing the human and societal aspects of the digitalization of our (industrial) workplaces, hence contributing to the human-centric perspective of Industry 5.0. Last but not least, projects look into the impact of digitalized work environment on workers' safety, working conditions, job satisfaction and physical and mental well-being (e.g. Human Manufacturing, SYMBIO-TIC, FIT4FoF, PLUS, MindBot, H-WORK, EMPOWER).

Industry 5.0 will be defined by a re-found and widened purposefulness, going beyond producing goods and services for profit. This wider purpose constitutes three core elements: human-centricity, sustainability and resilience. Rather than asking what we can do with new technology, we ask what the technology can do for us. It also means making sure the use of new technologies does not impinge on workers' fundamental rights, such as the right to privacy, autonomy and human dignity. Industry 5.0 recognises the power of industry to achieve societal goals beyond jobs and growth to become a resilient provider of prosperity, by making production respect the boundaries of our planet and placing the well-being of the industry worker at the centre of the production process. Technology serves people. The Factory2Fit project, for example, aims at empowering and engaging workers in a more connected industrial environment. The workers are given more influence and hence greater responsibility in shaping the production process, through virtual means. One of the fears associated with the uptake of new technologies is the loss of jobs. However, if applied correctly, new technologies have the potential to make workplaces more inclusive and safer for workers, as well as increase their job satisfaction and well-being. The skills dimension is another important set of considerations for Industry 5.0. Skills needs are evolving as fast as technologies. Digital skills are not the only skills that will be pertinent for industry workers in the factories of the future. The World Manufacturing Forum has identified a top-10 of skills that will be needed in future manufacturing. Surprisingly, only four of them refer to digital skills: "digital literacy, AI and data analytics," "working with new technologies," "cybersecurity", and "data-mindfulness". The remaining skills are more transversal skills linked to creative, entrepreneurial, flexible and open-minded thinking. The benefits for industry are wide-ranging, going from better talent attraction and retention, over energy savings, to increased general resilience. The overall benefit for European industry is longterm: continued competiveness and relevance by successfully adjusting to a changing world and new markets.

Summarizing, there are three main important aspects for the Industry 5.0. Firstly, talking about the human centricity, in order to ensure that both companies and workers benefit from the digital transition, rethinking and redesigning business models is necessary. Workers should be involved in every step of this transition process. In order to benefit from the relative strengths of technologies and workers, companies need to invest in both. Education, training, re-skilling and up-skilling are certainly among the most pressing issues to address when accommodating the digital transition in industries, as qualified human capital is of the utmost importance to make it a reality.

Then with respect to the sustainability, several powerful instruments helping the EU reach its carbon-neutral ambitions have been identified. Innovations in green technology, combined with EU initiatives aimed at Digitising European Industry (including better use of big data and artificial intelligence) are a reality and are increasingly embraced by industry. In the face of mounting public environmental and societal concerns, companies are incorporating sustainability into their business models. When fully realising the advantages of an improved corporate image and of savings on energy and material costs, industry will embrace resource efficiency as a natural choice.

Finally, With the Recovery and Resilience Facility, the European Commission wants to support EU countries in reform efforts that ensure sustainable recovery. Carrying out reforms and investing in the green, digital and social resilience priorities will help create jobs and sustainable growth, and allow recovery in a balanced, forward-looking and sustained manner.

As we have shown, the transition towards Industry 5.0 has already started. A number of on-going projects in Horizon 2020 are already contributing to the development of this concept. The following major actions, foreseen as next steps, are part of our growing toolbox for making Industry 5.0 happen:

- increasing awareness in industry,
- implementation of the technologies necessary for Industry 5.0. Its main outcomes are being taken into account in the preparation of the first Horizon Europe programme, in particular within Cluster 4,
- Identify existing actions and opportunities for the development of Industry 5.0 across Europe, including actions for encouraging inclusive technology diffusion across Europe.

2.4.2 Literature review

There are 88 articles in the scientific literature that mention Industry 5.0 in the title, of which 5 reviews [81] [82] [83] [84] [85] and 2 books, that are actually two volumes of the same book. [86] The main 4 works in this topic, according to the author, have been puplished from 2018 to 2021. [87] [88] [89] [90]

Investments in and expectations from Big Data have placed transnational research and implementation science communities under enormous and painful new pressures to rapidly edge toward innovative products and applications. Innovations, unprecedented by definition, do not necessarily follow a linear line from data to knowledge to application. They are important to remedy for a robust, sustainable, and responsible innovation ecosystem design in the digital age, and particularly for the networked large-scale scientific practices such as Industry 4.0. Industry 5.0 is as an evolutionary, incremental (but critically necessary) advancement that builds on the concept and practices of Industry 4.0. Others may wish to name it differently as "Industry 4.0 Plus", "Industry 4.0 Symmetrical", or "Industry 4.0-S". [87]

Although Industry 4.0 is not yet well grown, many industry pioneers and technology leaders are looking ahead to the Fifth Industrial Revolution: autonomous manufacturing with human intelligence in and on the loop. Its sole focus is to improve the efficiency of the process, and it thereby inadvertently ignores the human cost resulting from the optimization of processes. Unfortunately, Industry 4.0 does not have a strong focus on environmental protection, nor has it focused technologies to improve the environmental sustainability of the Earth, even though many different AI algorithms have been used to investigate from the perspective of sustainability in the last decade. Industry 5.0 will be a synergy between humans and autonomous machines. [88]

Concerning the Industry 5.0:

- humans are expected to add high-value tasks in manufacturing policies, where standardization and legalization will help to prevent any serious issues between technology, society, and businesses,
- particularly, senior members of a society and stakeholders will find it much more difficult to adapt with the new industrial revolution, [91]
- fast and highly efficient manufacturing may result in an overproduction phenomenon, therefore implementing transparency should also be taken into consideration,
- it is necessary to consider how autonomous systems can incorporate ethical principles,
- there should be explainable ethical behavior solutions in autonomous systems,
- ethical behavior in autonomous systems must be subject to verification and validation,
- essential skill gaps in future management and executive roles must be addressed.

According the literature, [89] [92] the following ones are issues related to integrating robots into organizations:

- evolution in organizational behaviour, structures and workflows,
- acceptance of robots in the workplace and social implications of human-robot co-working,
- evolution in work ethics, ethical issues resulting from human-robot co-working, and ethical status of robots,

- discrimination against robots or people,
- privacy and trust in a human-robot collaborative work environment,
- education and training, and learning to work with robots,
- redesign of workplaces for robots,
- legal and regulatory issues,
- psychological issues resulting from human-robot co-working,
- the changing role of HR and IT departments, and the emerging of robotics departments,
- negative attitude toward robots due to shrinking human workforce.

Summarizing another important work, [90] Industry 5.0 is currently conceptualized to leverage the unique creativity of human experts to collaborate with powerful, smart and accurate machinery. Many technical visionaries believe that Industry 5.0 will bring back the human touch to the manufacturing industry, [88] It is expected that Industry 5.0 merges the high speed and accurate machines and critical, cognitive thinking of humans. Another interesting benefit of Industry 5.0 is the provision of greener solutions compared to the existing industrial transformations, neither of which focuses on protecting the natural environment. [89] Industry 5.0 uses predictive analytics and operating intelligence to create models that aim at making more accurate and less unstable decisions. Despite a growing trend in Industry 5.0, we are not aware of any review article that focuses on Industry 5.0. Motivated by this observation, we aim to provide a very first review on Industry 5.0.

Definition 1. Industry 5.0 is a first industrial evolution led by the human based on the 6R (Recognize, Reconsider, Realize, Reduce, Reuse and Recycle) principles of industrial up-cycling, a systematic waste prevention technique and logistics efficiency design to valuate life standard, innovative creations and produce high-quality custom products. [93]

Definition 2. Industry 5.0 brings back the human workforce to the factory, where human and machine are paired to increase the process efficiency by utilizing the human brainpower and creativity through the integration of workflows with intelligent systems. [88]

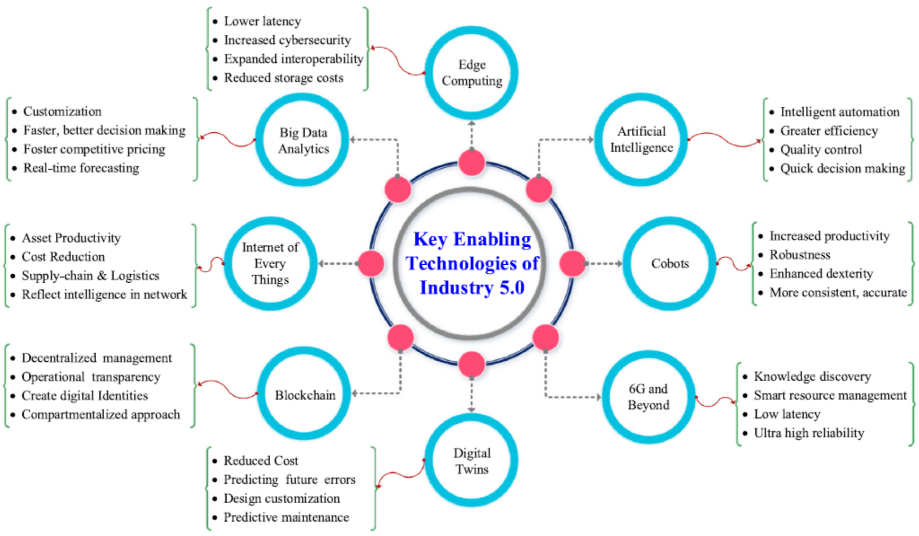

Fig. 2.13 Key enabling technologies of Industry 5.0. [90]

Definition 3. European Economic and Social committee states that the new revolutionary wave, Industry 5.0, integrates the swerving strengths of CPSs and human intelligence to create synergetic factories. [94] Furthermore, to address the human-power weakening by Industry 4.0, the policymakers are looking for innovative, ethical and human-centered design.

Definition 4. Industry 5.0 compels the various industry practitioners, information technologists and philosophers to focus on the consideration of human factors with the technologies in the industrial systems. [95]

Definition 5. Industry 5.0 is the age of social smart factory where cobots communicate with the humans. [96] The social smart factory uses enterprise social networks for enabling seamless communication between human and CPS components.

Definition 6. Industry 5.0, a symmetrical innovation and the nextgeneration global governance, is an incremental advancement of Industry 4.0 (asymmetrical innovation). It aims to design orthogonal safe exits by segregating the hyperconnected automation systems for manufacturing and production. [97]

Definition 7. Industry 5.0 is a human-centric design solution where the ideal human companion and cobots collaborate with human resources to enable personalizable autonomous manufacturing through enterprise social networks. This, in turn, enables human and machine to work hand in hand. Cobots are not programmable machines, but they can sense and understand the human presence. In this context, the cobots will be used for repetitive tasks and labor intensive work, whereas human will take care of customization and critical thinking (thinking out of the box).

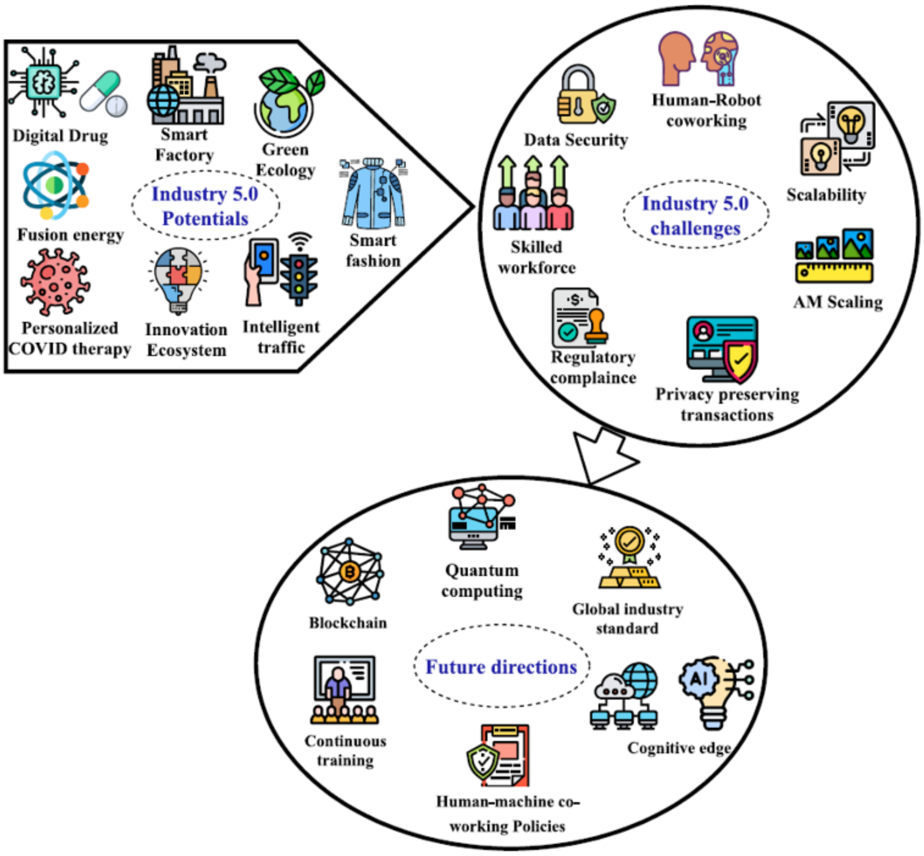

Fig. 2.14 Industry 4.0 potentials, challenges, and future directions [90].

Industry 5.0 is the enhanced version of the fourth industrial revolution. Another added features of Industry 5.0 are:

- Smart Additive Manufacturing (SAM) has become emerging technology in smart manufacturing domain,
- Predictive Maintenance,
- Hyper customization,
- Cyber Physical Cognitive Systems (CPCSs) with cognitive capabilities such as observe and study the environment and take actions accordingly these realizations, where learning and knowledge are the primary components of decision making that is also at the base of human-robot collaborative manufacturing.

Concluding, human intelligence can be applied for critical thinking of the customization logic, and the cobots can be utilized for labor-intensive tasks, thereby, alleviating the weakened human-power by effective use of cobots for labor-intensive jobs. Proactive Predictive Maintenance enables more manageable maintenance and a faster recovery rate in case of failures. The personalized manufacturing solutions through AI and cognitive systems for every customer will be assured by hyper customization throughout the manufacturing processes.

Chapter 3

Design framework for smart data-driven manufacturing services

The proposed framework has been structured to support the design phase of CPSs dedicated to the manufacturing sector in a Industry 5.0 context. In other words, the framework is a human thinking methodology to develop digital systems that actively operate as services family (assumed data-driven) with clear objectives and that require the awareness of the impact that their pursuing of goals has on the environment: digital services with arbitrary complexity that improve their sustainability level in the context of the CPS in which they operate.

Considering the works in the literature, the objective of this study is formalizing a graphical model to schematize the structure of a CPS for Industry 5.0, enabled by hybrid knowledge, i.e. hybrid modeling techniques for digital systems consisting of information management platforms, knowledge-based (cognitive) processes, and decision-support services with sustainable decision-support services (so operating with wisdom).

The framework has to guarantee scalable applicability: from simple services, e.g. a simple digital twin of a machine component (a more specific example is "the tool set actually on this milling machine"), to complex ones, e.g. an EIS that supports the human activity of an entire Industry 4.0 process. The framework is referred to a general IT solution that, being mainly data-driven, is highly complex to analyse, especially because it relies heavily on tangible and intangible assets with high-value and complex information baggage such as specialised machinery and the experience of a lean practitioner. This framework is designed to schematize the structure of any reactive data-driven services, i.e., systems that can not adequately be described by the relational or functional view, or otherwise with individual techniques, methodologies, or approaches. Typically, the main role of that systems is to maintain an interaction with their environment, and therefore must be described (and specified) in terms of their on-going behavior. This modeling approach supports the design of a generic cyber component: from a platform belonging to the classical EISs (e.g. a MES) to the digital twin of a simple product component (e.g. a mechanical properties simulation model). The framework must be able to be applied to any desired physical subsystem and to any human IT concept (with a minimum effort of generalization or specialization), but the application domain considered writing this work is given by systems of digital twins of products, material resources or machine components and other digital information management subsystems useful for human beings involved in a manufacturing process.

3.1 Basic concepts

This section reports definitions, fundamentals, and literature recalls in order to (i) fix the idea of applying the proposed framework as elementary structure for Cooperative Multi-Agent Systems (MASs), (ii) describe in the next section the proposed DIKW-structure of the agent, and, finally, (iii) discuss about the hybrid characterization of functionalities and properties of the agent for a sustainable, human-centric and resilient cognitive baggage.

3.1.1 Cooperative Multi-Agent Systems (MASs) modeling

The high-complexity level, given for example by the information load regarding projects and simulations during the design phase of a product, requires a framework that provides a manufacturing modelling approach of physical and (almost always) human-centric resources. In this scenario, the choice of assuming a basic information entity, a digital element (object), an item (a term widely used in PLM systems): an Agent-Based Model (ABM) is evident. Such agents can be associated with elementary productive resources without losing the property of being extended to a system of variable complexity (dependent on the estimated size of the set of agents largely influenced by it, e.g. as a mediator of communication). An agent is a reactive subsystem that exhibits some degree of autonomy, i.e., human beings has set objectives for it and the system that makes the agent works to achieve this task in the best way.

A multi-agent system (MAS or self-organized system) is a computerized system composed of multiple interacting intelligent agents. A multi-agent system is not always the same as an ABM. The goal of an ABM is to search for explanatory insight into the collective behavior of agents (which don't necessarily need to be "intelligent"), while MASs consist of agents and their environment. Typically multi-agent systems research refers to software agents. However, the agents in a multi-agent system could equally well be robots, human beings or human teams. A multi-agent system may contain combined human-machine teams.

The difference between a MAS and the object-oriented programming can be summarised by the following points [98]:

- objects are passive (no control over method invocation) while agents are autonomous (pro-active),
- objects are designed for a common goal while agents can have diverging goals (coming from different organizations),
- objects are typically integrated into a single thread while agents have own thread of control.

According basic concepts about MAS, an agent has the following characteristics [99]:

- **local view**, i.e. each agent has incomplete information or capabilities for solving the problem and, thus, has a limited viewpoint or the system is too complex for an agent to exploit such knowledge,
- **decentralization**, i.e., there is no system global control, or, in other words, no agent is designated as system controller or the system is effectively reduced to a monolithic system,
- **autonomy**, i.e., agents are at least partially independent, self-aware, autonomous and computation is asynchronous.

Other characteristics of agents are [100]:

- **proactiveness**, i.e., being able to exhibit goal-directed behaviour, or better if an agent has been delegated a particular goal, then we expect the agent to try to achieve this goal,
- **reactivity**, i.e, being responsive to changes in the environment,
- **social ability**, i.e., the ability to cooperate and coordinate activities with other agents, in order to accomplish their goals (as shown later, in order to realise this kind of social ability, it is useful to have agents that can communicate not just in terms of exchanging bytes or by invoking methods on one another, but that can communicate at the knowledge level).

3.1.2 Data, Information, Knowledge, and Wisdom (DIKW)

The framework is based on the characterization of the wisdom of a generic digital agent. The wisdom level is required for the agent in order to analyze its objectives and to generate solutions to pursue them. To ensure its sustainability, the agent has to generate smart decisions, i.e., decisions based on appropriate processes of continuous improvement (in a lean perspective) of the knowledge of the state of its universe, its self state, and all the space of its possible decisions. The DIKW pyramid, or scheme, is implied as general structure of the wisdom of a generic agent.

The main characterization [101] of works published with the aim of explore and clarify the possibility to define a hierarchical relationships between data, information, knowledge, and wisdom, is that:

- there is a clear consensus on the structure of the hierarchy and the definitions of the elements in the hierarchy,
- there is less consistency in the description of the processes that transform elements lower in the hierarchy into those above them, and some consequent lack of definitions clarity.

The research on Scopus for papers with DIKW (Data, Information, Knowledge, and Wisdom) in their title produced 32 results at the end of 2021, of which 12 journal articles, 15 conference papers, 3 book chapters, and 2 reviews shown by Figure 3.1. These works don't include all the previous works carried out on the same theme but which do not use the term "DIKW".

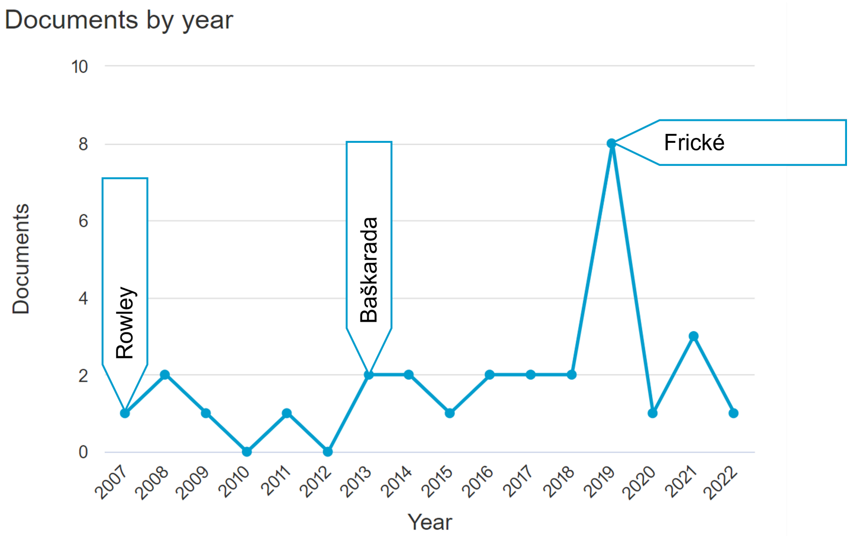

Fig. 3.1 Number of documents with the therm DIKW in their title. [101] [102] [103]

The conference papers are only minor works with few citations, while, between the 12, the main three journal articles talk about:

- a first representation under the therm "DIKW" of a wisdom hierarchy that is the most cited one from the 2007, [102]
- a 2-years later critique of this hierarchical representation [101] that is the second most cited work on this topic,
- a further contribution focused on the quality dimension of the hierarchy, [103] published in 2013.

Between the three books, the most inherent one exposes a methodology for a purpose computation-oriented modeling based on the DIKW architecture [104] and explain how in the higher level (the one dedicated to the wisdom) has to express the purpose of the system modeled by a DIKW structure. Between the two reviews, instead, the interesting one is written by the same author that ten years before analyzes few critical points of the DIKW structure. [105]

In more recent literature, authors often cite Ackoff's paper of 1989 as a source for the hierarchy. [106] This article proposed a hierarchy with the following levels:

data, information, knowledge, understanding and wisdom. It included understanding in his hierarchy, but more recent commentators have disputed that understanding is a separate level. Ackoff offers the following definitions of data, information, knowledge and wisdom, and their associated transformation processes:

- **data** are defined as symbols that represent properties of objects, events and their environment and they are the products of observation but generally in a not usable state (i.e. relevant) form (the difference between data and information is functional, not structural),
- **information** is inferred from data and contained in descriptions, answers to questions that begin with such words as who, what, when and how many, while information systems generate, store, retrieve and process data,
- **knowledge** is know-how, and is what makes possible the transformation of information into instructions and it can be obtained either by transmission from another who has it, by instruction, or by extracting it from experience,
- **intelligence** is the ability to increase efficiency,
- **wisdom** is the ability to increase effectiveness and it adds value, which requires the mental function that we call judgement (the ethical and aesthetic values that this implies are inherent to the actor and are unique and personal).

Considering that there is more data than information, than knowledge, than wisdom, [102], various works hypothesis different views on the variables that change between the different levels of the hierarchy: low levels (data) are associated to an high perception of computer inputs and programmability or algorithmicity (automation), while high levels of the pyramid (wisdom) are associated to an high perception of meaning, value, human input, structure and human agency, transferability, actionability, and applicability.

Data has to be more than the mere "observable", and it can be more than the pronouncements of "instruments". There are contexts, conventions, and pragmatics at work. In particular circumstances, researchers might regard some recordings as data which report matters that are neither observable nor determinable by instrument. An interesting lesson in the literature for information studies is: "do not suppose that there is a special category of "data" which can serve as the bedrock for all else." [101]

Information is notoriously a polymorphic phenomenon and a polysemantic concept so, as an explicandum, it can be associated with several explanations, depending on the level of abstraction adopted and the cluster of requirements and desiderata orientating a theory. [107]

The essence of information science is that it deals with records, recordings, documents, inscriptions, and representational artefacts. Its historical origins are librarianship, archival studies, and the theories and practices of documentation. Nowadays, research in information science has spread widely from its historical base but its core is still the attention to those artefacts of preservation of forms of bridges connecting a individual and instant of time with availability across individuals and persistence through time. There are even many different senses of "information" in use even in information science and, clearly, it is not the case that one of these senses is good (and all purpose) and the others not. But, both in information science and elsewhere, there are different problems and different contexts where these different notions of information come into play.

Information systems books tend to focus on the relationship between data and information, often defining information in terms of data: "information is formatted data", "information is data that have been shaped into a form that is meaningful and useful to human beings" or "information is data that have been organized so that they have meaning and value to the recipient". [102] The word "information" has been given different meanings by various writers in the general field of information theory. It is likely that at least a number of these will prove sufficiently useful in certain applications to deserve further study and permanent recognition. It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field. [108]

So much for data and information in the DIKW hierarchy, the pyramid has no solid foundations for the knowledge. Within philosophy, there is the distinction between **"know-how"** and **"know-that"**. The know-that is always in a propositional form and, given a suitable expressing language, they can be written down and recorded or stored in data-bases. Know-hows are different: they might be articulated as procedural rules, usually ‘if–then’ rules or more general regular expressions, but not always. Knowing how to solve a quadratic equation, how to control a parameter, and similar, might be conceived like this. Such rules, of course, can be written down and stored in a repository. Other know-hows do not seem to be of this kind. Knowing how to ride a bicycle (or satisfy a customer) is not plausibly a matter of the brain scanning, and selecting among, rules like "if you want to turn left, lean left" [101] (or "if a customer needs a product, produce it and sell it at in the most sustainable way"). Another common assumption is that certain knowledge does not exist: all knowledge is conjectural. Such scenario, originally called fallibilism, considers that propositions concerning empirical knowledge can be accepted even though they cannot be proven with certainty, or in short, that no beliefs are certain. [109]

3.1.3 State-of-the-art about hybrid modeling

CPSs share mathematical characteristics too, which are in many ways more important for the aim of this framework than the fact that they happen to be built from cyber components and from physical components. From a mathematical perspective, CPSs are hybrid systems (or extensions thereof). Overall, hybrid systems are not the same as CPSs. Hybrid systems are mathematical models of complex (often physical) systems, while CPSs are defined by their technical characteristics.

Hybrid modeling is the process of making use of two or more modeling technique belonging to different philosophies or methodologies and then synthesizing the results into a single score or spread. The main works on this topic use the following therms (with the following frequencies): "hybrid modeling" (37), "grey/gray-box" (11), "knowledge based modular networks" (1), "incorporating external information" (3), and "semi-mechanistic model structures" (1). [17] Hybrid modeling is conceptually similar to ensemble modeling but, while the seconds are referred to unify different models of the same family, hybrid modeling is referred to completely different approaches of modeling with a consequent increasing in therm of complexity of managing heterogeneous characteristics.

It is clear, therefore, that this research topic is widely covered considering the different terms by which the scientific community refer to it. This work aims to provide a proposed definition of this approach, answering the second research question (RQ2), by formalizing the characteristic of hybrid at the level of data, information, knowledge and wisdom, considering that the definition of hybrid knowledge, i.e., hybrid KBS, turns out to be the most dealt with, even in the case study, because it is considered the most interesting by the scientific community.

A literature research executed on *Scopus* is the base of the discussion regarding the origin of the therm, and more specifically the first research query is shown by the Figure 3.2: the aim is finding all the English-written works in the area of math, engineering, and computer science, that are related to the concept of hybrid modeling.

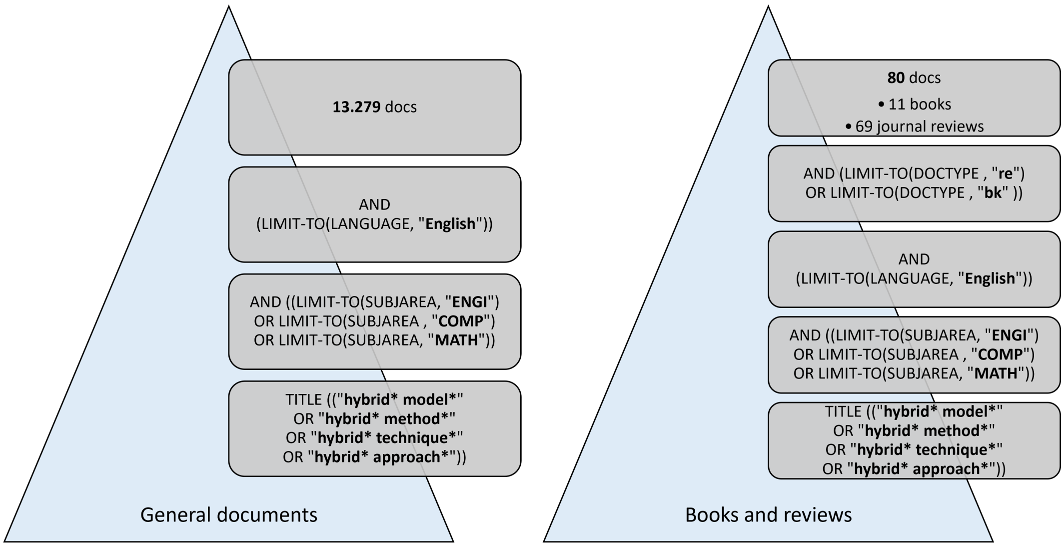

Fig. 3.2 Scopus research: all the English-written works in the area of math, engineering, and computer science, that are related to the concept of hybrid modeling.

The Figure 3.3 shows the number of publications in the years starting from the publication in the 1962 [110]. This first work starts talking about the benefits given by the use of a combination of both analog and digital computing devices: the conclusion is that the addition of the digital expansion system providing logic capability to the general purpose analog computer will enlarge its capabilities as an automatic computing device.

Another work published in the 1965 uses the therm hybrid to identify computers that are both digital and analog and it highlights that the term had already been in use for some years and had originated in the aerospace industry to refer to the technology behind the computers used for simulations. [111] Two years later, in the 1967, the therm hybrid was used to indicates an approach to solve non linear equation based on two main techniques: the quasi-linearization one and the adjoint state method. [112] While the quasi-linearization technique is essentially a generalized Newton–Raphson method for functional equations, the adjoint state method is a numerical method for efficiently computing the gradient of a function or operator

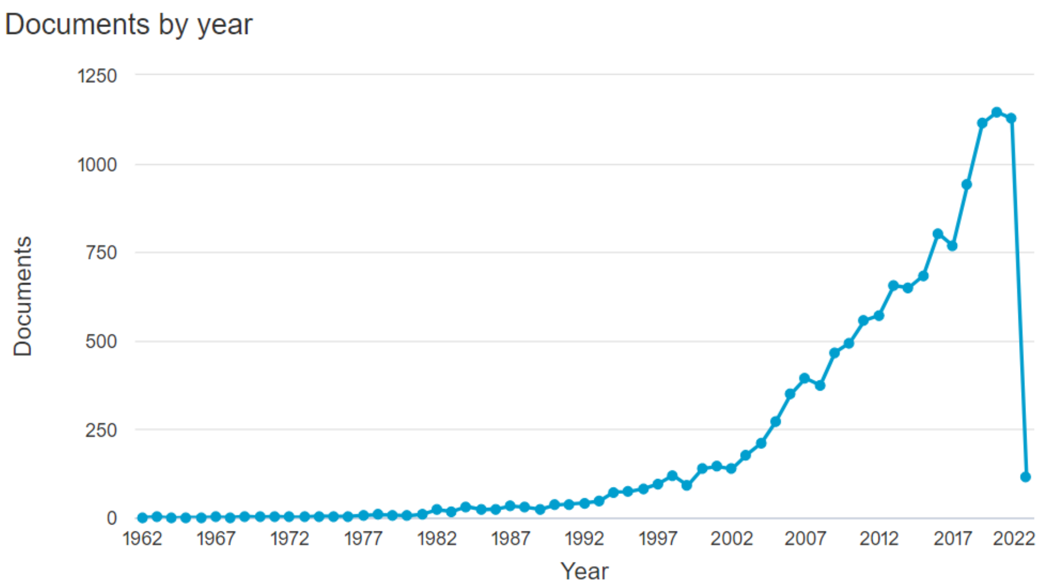

Fig. 3.3 Scopus research: all the English-written works in the area of math, engineering, and computer science, that are related to the concept of hybrid modeling.

in a numerical optimization problem. The work shows that the hybrid use of the two methods greatly reduces the number of first-order differential equations to be numerically integrated. A further work published in 1969 uses the therm hybrid for a root-finding method that is inspired by a previous work of the 1961 and makes use of analytical techniques, that is statistical in nature, in combination with short tables, i.e., a deterministic method. [113]

The hybrid methods consider models of different nature to take advantage from the synergy in particular of the following combinations:

- analog and digital components,
- analytical equations and numerical methods,
- statistical and deterministic approach.

"The diversity of backgrounds led to different terminologies for describing the combination of mechanistic and data-driven models (hybrid, gray-box, etc.). Furthermore, different understandings exist of what is really an hybrid modeling approach; from broader definitions, where data may just be used to adjust mechanistic parameters, to stricter definitions, where two dissimilar sub-models (mechanistic and data-driven) are required to be simultaneously present." [17]

In conclusion, the term hybridisation, which refers to a type of analysis modelling, originated in the 1960s in response to the need to overcome the complexity limits of classical models by taking advantage of new computational capacity. The term refers to a main analysis technique that makes use of several techniques of different natures in order to determine the optimal solution to a particular problem of such complexity that classical analysis techniques are ineffective and the pure application of numerical methods is computationally too costly.

In the era of Industry 4.0, internet of things combined with generalized advanced digitalization, business analytics and additive manufacturing will drive industry to a market more decentralized, flexible and customizable, focused on client-by-client solutions and with reduced time- to-market due to rapid disruptive innovation. With the exponential increase of the amount of data available, hybrid modeling requires new frameworks and tools for combining first-principles with data-induced knowledge, including for instance the ability to incorporate new types of data and for dealing with the 5 Vs of Big Data. [17]

The **spectrum of modeling techniques** used by the literature are the mechanistic modeling and the data-driven one. The construction of a **mechanistic model** for hybrid modelling frameworks, depends on the available prior knowledge. These mathematical statements can be expressed more simply as algebraic equations or, with increasing complexity, as ordinary differential equations (ODEs) (for lumped parameter system modeling), differential algebraic equations (DAEs) or PDEs (for distributed parameter system modeling). By increasing the complexity of the mechanistic model, one reduces the structural mismatch between the model and analyzed process, but the larger number of parameters can lead to an unsuitable model if there is not a way to reasonably estimate said parameters (this is the bias-variance trade-off for mechanistic models). In other cases found in the literature, process knowledge is encoded in some equations that are not available and this does not mean the absence of such knowledge: process flow diagrams and even engineering expertise can provide enough information to improve a data-driven model. The construction of a **data-driven model** is argument of data science and applied, for example, in nowadays deep learning applications. It is imperative to recognize the importance of data-driven models in a hybrid structure, specially in the increasingly data intensive industrial settings. The first wave of scientific papers on hybrid modeling, mainly used data-driven techniques to fill in the gaps of knowledge in mechanistic models, such as unknown nonlinear behavior (e.g., kinetics) or unknown parameters suspected to have a complex dependency on the process variables behavior. Currently, as data-driven modeling becomes more widely accepted as a legitimate and useful technology for analysing process data, an opportunity arises to apply them not just as a complement to mechanistic models, but as a way to model which may benefit from the incorporation of prior knowledge and process understanding.

Finding the right **hybrid model structure** is fundamental. In the classical hybrid modelling literature, the prevalent way to combine the two modeling approaches (mechanistic and data-driven) starts from an analysis of the structure of mechanistic model and its assumptions. A parallel configuration can compensate for mechanistic structural mismatch, but if the mechanistic structure is accurate enough, then a serial configuration is usually a better choice. [17] In the **serial structure** approach, the "white box" (mechanistic model) and the "black box" (data driven method) are combined in such a way that one provides an input for the other (use a neural network to estimate a parameter or use a mechanical equation to calculate a feature for inference). In the **parallel structure** approach, usually the mechanistic prediction power limited due to limitation in describing some effects is improved by the data-driven model. The parallel structure, despite being adequate for handling model mismatch, still relies on a robust mechanistic model. An example is the coupling a data-driven model, comprising a time domain partitioning procedure, with a mechanistic one: the data-driven model delimits zones where the mechanical model can well describes residuals. The hybrid approach often outperforms purely mechanistic models, but its superiority cannot be guaranteed against purely empirical models particularly if the mechanistic model does not make a good representation of the physical behavior of the process. **Surrogate models** (or substitute models, or meta-models, or response surface models) are simpler mathematical representations of more complex models. They require less computational effort to be run than the more rigorous representations, and have been extensively used in process modeling and optimization. These models are designed to yield unbiased predictions of sampled or simulated data which is useful to generate regular measurements in complex systems. Another way to design a surrogate model is to generate data points from a complex mechanistic model and to use them to train a data-driven model (estimate parameters for simulating a complex process and make subsequent analysis).

Once the structure of the hybrid model is defined, it is necessary to estimate the models' parameters (**hybrid model or machine training**) solving an optimization problem. In the literature about hybrid structuring, usually, priority is given to the mechanistic model and once this model is set, the data-driven model is identified with standard techniques. Besides this direct approach (or better **mechanistic-first approach**, researchers also proposed an **incremental approach** (decomposition of large problems into smaller ones trough decomposition algorithm), a **sensitivity approach** (training the hybrid model by back-propagating the errors through the ANN and the associated mechanistic model) and an **evolutionary computing approach**.

Like in CPS or IoT systems, multiple simple components combine for a common bigger cause and their combination compounds their interaction and impact generating new challenges , similarly, despite the significant recent progresses, there are various challenges regarding hybrid modeling [1]:

- there are no clear demonstrations about benefits given by hybrid modeling and no clear characterization of processes that require it,
- there is no a common definition of this technique,
- there are no model selection clear guidelines,
- there are only few comparisons between hybrid models and single ones,
- there are no enough benchmark problems and relative dataset for evaluations and comparisons,
- prior knowledge, although dominated by equations, has not been fully incorporated in all its configurations,
- data-driven models have been used to improve mechanistic models, but the opposite path remains vastly under-explored,
- there is a need for the development of software tools that facilitate the incorporation of the various sources of knowledge.

3.2 Wisdom characterization of the agent

This framework is based on a DIKW structure used to model any computer programs, clearly not human cognitive processes, achieving expert-level competence in solving problems for manufacturing task areas. Considering a software application working on a CPS, the DIKW structure can be used to model the entire one and all its sub-components (often dedicated to specific operative tasks or physical resources) focusing on characterize data, information, knowledge, and wisdom involved in the process and how model them to design a smart agent. Any agent, operating in a system, is a system itself characterized by a DIKW structure. The agent of a CPS, therefore, represents any entity acting, actively or passively, in that system. First, in any design of a digital platform it is necessary to model the CPS that is to be developed, and this represents an initial agent representing the entire system operating in a certain environment with its KPIs set and its vision. Subsequently, in the design of a manufacturing CPS, it is necessary to consider some conventional systems such as ERP systems, PLM systems and MESs as agents operating in the information system: in fact, it results spontaneously to call them systems itself or sub-systems rather than agents. The different components or modules or functionalities of these software, if necessary to model them, consist of a wide variety of different types of agents: SCADA systems, gateways and, for example, all other entities defined by the ANSI/ISA-95 standard. [114][115]

Therefore, modeling a CPS means modeling different physical manufacturing resources (and their interaction) such as humans, machinery and tools, products, and resources: in other words, just as it is necessary to create agents representing conceptual functions or single functions of physical resources (scheduling for example is a single function executable by a manufacturing engineer), it is necessary to create more complex agents represented entirely by a physical resource (e.g., a prototype or a specific milling present in the factory). In this case, agents referring to such resources can be considered digital twins. For these cases, this work aims to provide a definition of digital twin that is not limited to using simulation methods to know the various states of the agent, but makes use of hybrid modeling based on different modeling methods.

Considering the case of agents referred to human beings, each agent can be considered a weak AI itself, as opposed to the strong one called Artificial General Intelligence (AGI), [116] [117] and, therefore, is referred to an agent endowed wit the functions that the human has in the manufacturing context. Each agent is an AI not intended to replicate human decision-making processes, but with the aim of providing a model underlying agents capable of replicating various human functionalities (knowing the work status of an operator, scheduling the work to be done in the day, or improving the ability to assemble two particular components). It is important to emphasize that limiting the model to a description of the human from a manufacturing point of view does not mean not considering its complexity; in fact, as explained in the following sections, such agents play a key role in factory wisdom and for this reason providing a simplified model of them is rarely sufficient: for example, designing any type of software, including humans as agents of the model with strong requirements in terms of Human-Machine Interface will surely ensure to produce a CPS closer to the goals of 5.0 era.

3.2.1 DIKW-MAS: a system of agents with a holistic wisdom

This work starts considering an agent that is a purposeful producer of actions, that requires knowledge referred to the associated physical resources, or more generally, that has an appropriate data structure to extract the knowledge base. In other words, the data generated along the life cycle of a CPS can be imagined as data set organized and managed by a set of digital agents that live in a digital system. Summarizing, a MAS can be considered more general than ABM and is based on the concept of an heterogeneous agent that lives in its environment, and interacts any agent of the same environment actively, passively, and as a mediator (middleware).

Therefore, this work proposes a multi-agent model that is based on the following hypothesis:

- every digital object can be considered as an agent (including the simplest agent represents by a "single datum" and the most complex one, if one exists, called "ambient system"),
- every agent refers to **data** bases to which it has access,
- every agent uses the **information** given by receiving, linking, managing, and process data (considered as the data management and mining level),
- every agent has a **knowledge**, that is a set of models, rules, logic expressions and human protocols that make the system intelligent, or rather, capable of understanding and using the information (the level that enable decisions, whether they are awareness human decisions due to insights generated by the supporting machine or decisions that the human has relegated to the same DSS),

- every agent has a **wisdom**, that is a set of objective functions that make the agent a smart, i.e., a decision-making, component operating in a CPS with knowledge of itself and other agents according to distributed interaction rules.

The complete autonomy is an unquestionable characteristic. The local view is guaranteed by the limited information managed by a single agent according its objective of maintain a sustainable existence: in a 5.0 agile context, even the most complex information systems are designed according to a structure based on synergistic components that can be modelled as agents. [118] [119] In such framework the agent has a view limited to a certain number of other agents and their information and this limit could be change to improve the agent performance but it definitely remain confined to a relatively low number of agents.

The concept of decentralization is respected, i.e. the interaction between agents is controlled by the common system, but it is extended considering that each agent can act as system and therefore an agent can always be considered as a system. Concluding, an agent can belong to one or more agents that act as systems for It and, at the same time, the same agent can act as system for one or more agents. This guarantees the hybrid being of the model that is in the middle between a MAS and a monolithic one, in order to limit the design complexity of a hypothetical overall industrial information system.

Two features of multi-agent learning which merit its study as a field separate from ordinary machine learning. [120] First, because multi-agent learning deals with problem domains involving multiple agents, the search space involved can be unusually large; and due to the interaction of those agents, small changes in learned behaviors can often result in unpredictable changes in the resulting macro-level ("emergent") properties of the multi-agent group as a whole. Second, multi-agent learning may involve multiple learners, each learning and adapting in the context of others; this introduces game-theoretic issues to the learning process which are not yet fully understood.

As resumed in Figure 3.4, the general proposed DIKW structure includes:

- **4 hierarchical levels** (Data, Information, Knowledge, Wisdom), where data level is not properly owned by the agent but it is more a active window on which the agent can operates trough the information level,

- **description of the top-down and bottom-up functional flow** , i.e., in Figure 3.4 there are the names of the contribution that a level of the DIKW-pyramid gives to the upper and lower one,

- **7 axis to characterized the DIKW-hierarchy**, where according 5 the wisdom is in the higher part of the scale, while following the other 3 is in the minimum one.

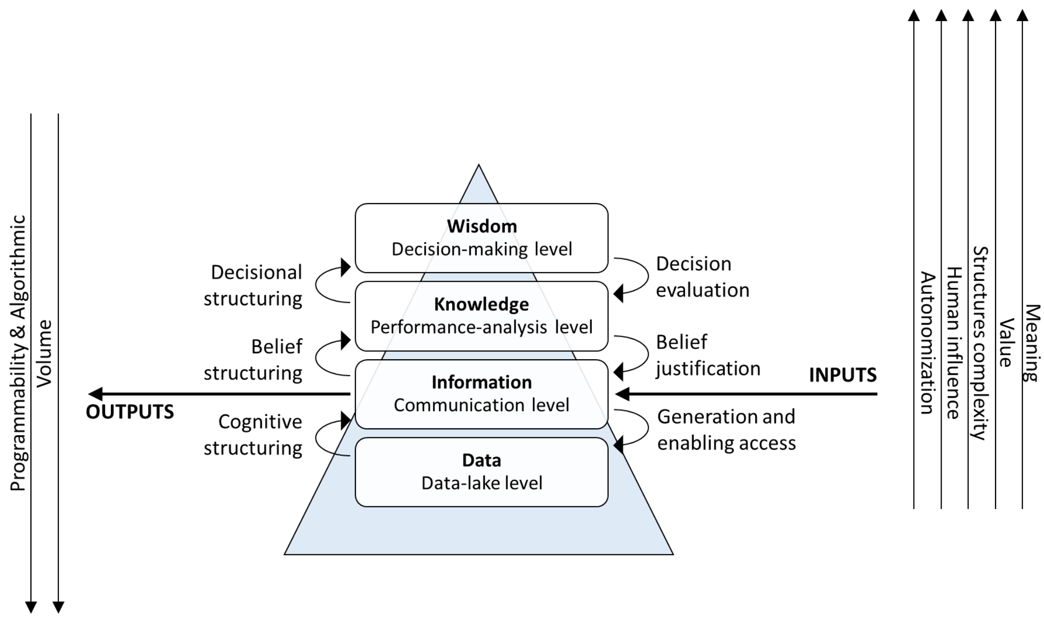

Fig. 3.4 DIKW structure of the agent (the system).

The agent is assumed a wisdom-based system, although, according to the definition, there are cases where the agent, for its simple objectives, can be considered as a simpler systems without few levels of the DIKW-structure, with a hierarchical constrain: to design each level of the pyramid, each level below has to be designed. In other words, agent

- every agent is a data base (DB) system considering that the simplest one has the same complexity value of the string "Hello world!",

- for every agent, the information level is schematize to design the information relationships which include the data of the agent,

- (if needed) for every agent, to schematize the knowledge level the agent has to be an information system, i.e., an agent with a structured information level,

- (if needed) for every agent, to schematize the wisdom level the agent has to be a knowledge-based system, i.e., an agent with a structured knowledge level.

Always considering the assumption of equivalence between agent and system, thus between an information agent and an information system, few examples of information agent are a standard ".csv" file, i.e., i.e., an agent required a tabular information structure, and a DBMS without the knowledge level but presenting a deep and technical treatment of the information one. Otherwise, a Tool Condition Monitoring (TCM) system dedicated to a drilling tool is a KBS without the wisdom level but that probably presents the formalization of predictive models to estimate the status of the tool based on production and sensor data. An agent can be a human in the cyber physical system interacting along the cyber-physical interface (a machine operators, a production manager or engineer, a maintenance specialist, or a product designer). Further examples of agent can be a digital service, one from all, a digital twin for a product (or family of product) or for a stockable manufacturing resource.

The framework consider the following definition for the DIKW-pyramid:

- **data** represents simply the "data lake" of all available data to the agent trough its information level,
- **information** is about technologies for generating, managing, and share data with other agents (why not human),
- **knowledge** is information (always data) that express expert opinion, skills, and experience, to result in a valuable asset which can be used to aid decision making (by any types of other agents like human),
- **wisdom** is the way of evaluate the best choice considering a certain knowledge.

The 7 views on the variables that change between the different levels of the hierarchy are considered according the following definitions:

1. **meaning** (↑) axe refers to the contribution that the level gives to the smart or intelligent being of the agent,
2. high **value** (↑) is referred to smart data, i.e., less data generated by the wisdom can impact drastically more respect the same quantity of generic data,

3. **structures complexity** (↑) is higher for the wisdom because physically expressed by less data with more meaning,

4. **human influence** (↑) is related to how much the human can/has to decide how processes in a level work,

5. **autonomization** (↑) is something far from simple data that alone are almost always not available,

6. **programmability** (↓) is more accessible (higher values) in therm of mechanistic explications,

7. **algorithmic** (↓) refers to levels where it is simple automatize processes;

where in brackets it is specify the direction of the axe with an upward arrow (↑) or downward arrow (↓) according Figure 3.4.

The following processes (can be consider as inter-level functions of the agents) generate wisdom from data:

- **cognitive structuring** (data → information) of data is required for information,
- **belief structuring** (information → knowledge) is the structure to give meaning to (select, synthesize, and understand) the information that is significant "to know",
- **decisional structuring** (knowledge → wisdom) is the schematizing of the knowledge required to evaluate decisions, to generate insights.

Following the opposite direction, the following processes express a sort of backpropagation flow on which the wisdom is based:

- **decision evaluation** (wisdom → knowledge) consists in considering knowledge, taking decisions and evaluating them (know them),
- **belief justification** (knowledge → information) is the continuum process of knowledge "uploading" trough new information,
- **generation and enabling access** (information → data) make possible digitization processes, data mining activity, and data shared systems.

Finally, the interface with the environment is managed as a flow of inputs managed in the information level, as shown by the Figure 3.4, dissipated in other levels if needed, and returns to the environment (or to systems to which it belongs) as outputs: "one data flow" may change the state of the agent, whose output flow will certainly influence other systems as it has been asked.

3.2.2 D: data lake

Data is the result of any digitization activity. This framework consider data generated along the interaction between the software and the final users for which the software is designed. However, it is reasonable consider that an amount of data are physical associated to an agent (e.g., if a smartphone is modeled as an agent its memory can be data for other agents, but physically they are link to the agent smartphone: the smartphone is the only agent with the maximum level of property respect such data).

From the birth until all its life cycle, an agent is characterized by data under its view and it can perform activities to pursue its scopes. Data are considered having the following characteristics [102]:

- data has no meaning or value because it is without context and interpretation,
- data are discrete, objective facts or observations, which are unorganized and unprocessed, and do not convey any specific meaning,
- data items are an elementary and recorded description of things, events, activities and transactions.

The acquisition of data can be generalized well beyond automatic instruments. When, for example, a person fills in a form giving their name, address, age, social security number, these inscriptions are data (actually, the term "raw data" seems apposite). [101] This instruments to acquire, store, and retrieve data are arguments of the information level of the DIKW-pyramid. Following this idea, in the proposed framework data are not property of the agent, but they are only selecting by them; in other words, it is clear that part of data are stored physically somewhere but it is not obvious to define, for example, which Data-Base Management Systems are managing such data.

Concluding, saying that the data level has the highest portability is true but it is obviously the most limited by physical constrains. The Industry 4.0 has given the freedom of thinking systems hyper-connected, almost with no limits of accessibility, but it is clear the vision given by an 5.0 point of view where "connections", or data-exchange edges, are available if sustainable first and foremost for all humans but also for the entire environment around us and in all conditions, even the most critical. This limited view is a peculiar characteristic of each agent that decide with a certain wisdom of selecting its data considered smart, i.e., a sub set of smart data selected from the data-ocean that is technologically available for every system in the same environment.

3.2.3 I: Enterprise Information management Systems (EISs)

The vision is that of a human asking a question beginning with, perhaps, "who", "what", "where", "when", or "how many" [106] and of row data processed as an answer to become information: data itself has no value until it is transformed into a relevant form. In consequence, the difference between data and information is functional, not structural. Information can also be inferred from data (it does not have to be immediately available). Talking about information as the result of processing, interpreting, classifying, rearranging or sorting, aggregating, performing calculations, and selection, it is clear the connection between these processes and ones involved for example in activities like data mining or feature engineering.

Such processing of data requires a decision about the type of analysis, and this, in turn, requires an interpretation of the content of the data. To be relevant and have a purpose, information must be considered within the context where it is received and used, i.e., being the meaning a subjective notion different agents consider valuable different data and are in this way different types of information systems. [121] Every agent (or family of agents) has an information level structure that has view on a set of data and make them available for itself, e.g. for decision making, or for other agents (for example an agent representing a human being or an ERP system already involved in the same process). The agent can be considered as an information management system that, beyond the concepts of knowledge and wisdom, makes accessible the information contained within the database over which this IT agent has jurisdiction (physically or trough the network).

Examples of information agents are: (i) a digital twin is an information system clearly developed to generate a specific knowledge about a product or a manufacturing resource, (ii) common Data-Base Management Systems (BDMSs) can be considered as the most complex information system without specific knowledge to be fulfilled, (iii) the information structure of an ERP system is commonly given by a functional division of resource planning activities with the aim of developing specific knowledge (several ERP solutions on the market consider different business functions as stand-alone information modules, agents, able to exchange information between them and aimed at generating specific knowledge such as about customers, suppliers, warehouse or other functional areas of the company), finally, (iv) a MES in a Industry 4.0 scenario has an information level probably high depended by the software interface with manufacturing plant machines and consider such machines as agents or family of agents with its own information level.

Concluding the information level of the framework is required to describe all the IT services and micro-services used by the agent to generate, maintain, and share information regarding the context in which the agents operates in form of data, i.e., this is the levels of all modeling technique finalized to make efficient (sometimes autonomous, sometimes automated, sometimes simply organized) the information flow between digital and physical resources and the resulting structured increase of the database to which agents have access. It is important to underline a common view in literature where the information scientist does not want to be collecting data without being sure that they are promoted to accessible information (science like cyber-security and digital privacy investigate about the data accessibility).

A better data-acquisition methodology is more top-down and just-in-time: data are generated trough processes in the information levels and it is trough this level that a human can receive exactly the data-driven (digital) information needed to answer a particular need. In this level there are information structures that change radically among systems in the same environment (just think of the difference between two systems in the same manufacturing environment such as the ERP platform and a service of edge-computing of a fire alarm system installed in one single smart sensor. Complexity, Human-Machine interfaces, strong-weak belonging to a system, and other peculiar agent characteristics make this level the least portable of all: the portability can be considered as limited among agents of the same family, i.e. agents that have the exact same information structure but differ due the historical information regarding experiences and choices (e.g. the software solutions installed in different smart fire-alarm sensors around the plant).

3.2.4 K: Knowledge-Based Systems (KBSs)

Knowledge, in the sense of a knowledge base or knowledge within traditional philosophy, is assumed just as a collection of "know-thats". When it has, for example, an high positive impact on human decisions, this "know-that" is common defined "insight". [122] In this framework, according to the inter-level processes highlighted in the Figure 3.4, the therm knowledge level refers to models for the "belief" structure, i.e., a sort of "what it is possible to know", models for estimating all "know-thats" and their affidability levels, i.e., answering to the question "how much is true this know-that", and finally models that estimate new types of knowledge, or better that is able to know that there is a "new" knowledge not included in the knowledge structure of the agent that is valuable for the agent, vaguely similar to wondering why something is unknown.

Including all inter-level processes involved, the knowledge level is characterized as following:

- the **belief structuring** (information $\rightarrow$ knowledge) is the model used to select relevant information describing the agent (significant variables, predictors or KPIs),
- the **belief justification** (information $\leftarrow$ knowledge) is the management of the knowledge-base mined by the information level,
- the **decisional structuring** (knowledge $\rightarrow$ wisdom) is the model of solutions that the knowledge level uses with the wisdom one,
- **Decisions evaluation** (knowledge $\leftarrow$ wisdom) allows the wisdom level to evaluate the environment state referring to a single its decision

Thus, a knowledge-based system has two distinguishing features: a knowledge base and an inference engine. The first part, the knowledge base, represents facts about the environment and the agents itself. The second part, the inference engine, allows new knowledge to be inferred (like the knowledge about the need of new knowledge). For these characteristics the knowledge is more portable than the information one, however, it remains linked to groups of families of agents linked by a common information level main characteristics and it works like a bridge between the least portable level, the information one, and the most portable one, i.e., the wisdom.

Concluding the definition of such level, the knowledge component of an agent can be assumed as a KPIs set that describes different states of the referring agent (real-time states o forecasted ones, like for a digital-twin system), using and generating information.

3.2.5 W: Decision-Support Systems (DSSs)

Considering the wisdom as the level referring to the decision-making activity, wisdom can be considered the ability to use the knowledge regarding the system in order to find the best decision respect the decisional structure. Data are the physical base of decisions, information is the instrument to use this base, and finally the knowledge is near to what is considered to make decisions. Trough its wisdom an agent pursues its goals depending by the ones of the environment, and trough its data the system receive knowledge regarding the agent states and actions. Everything starts from a scope.

Regarding DIKW-structures for EISs, the main lack in the literature is a treatment of the wisdom that makes it usable in conceptual applications like this framework. It is clear that wisdom is a very elusive concept that has more to do with human intuition, understanding, interpretation and actions, than with systems. [123] Some interesting works assert that wisdom is the ability to act critically or practically in any given situation, it is connected to the use of information, knowledge, and ethical judgements related to an individual's belief system: an accumulated knowledge, which allows you to understand how to apply concepts from one domain to new situations or problems. [123] [124] [125]

Wisdom therefore has to respect more than any other level the requirement of portability (among agents): it has to be a model which tends to be standard in the environment which characterises the operation of all systems included in that environment. This characteristic, referring to a component of the agent dedicated to decisions, can be associated to the common vision that humans of the same organisation should have: a vision that "suits" all the members of the organisation (a manufacturing plant, a company, a consortium or a commercial network).

According to the inter-level processes highlighted in the Figure 3.4, wisdom is the level responsible for the agent's decision-making structure and the action of making a decision and extracting knowledge about its impact:

- the **decision structuring** is given according human philosophy and culture (manufacturing examples are PLM, lean thinking or JIT, and Industry 5.0) and it gives guidelines about required knowledge,
- the **decision making** is possible trough optimization models (e.g. minimizing, ranking or selecting) that return a digital representation of the best solution giving to a problem,
- the **decision evaluation** processes the knowledge representation of the decision-made to evaluate the impact of the decision for optimizing the wisdom level (systems of models).

In this sense, the wisdom is considered as a decision-making management level that is portable among different systems (probably also among different environments or fields) and it communicates with the knowledge system delivering solutions to its problem formulations (one solution per time for each), considering (knowing) the history, the actual state (digital twin), predictions about future, and obviously the solutions space of the knowledge level. Referring to the axis in Figure 3.4, this level has the greatest values of meaning and structures complexity because referred to complex events systems, value in order to obtain a sustainable manufacturing system, human influence for its aim of being a formalization of human needs, autonomization powered by AI technologies.

3.3 HW-MAS: Hybrid-Wisdom characterization of the agent

According to the literature, the DIKW theory is poorly covered, well structured, but considered an uninspired methodology by [101]. In this framework, the DIKW theory is described in detail according the author assumptions and structured to

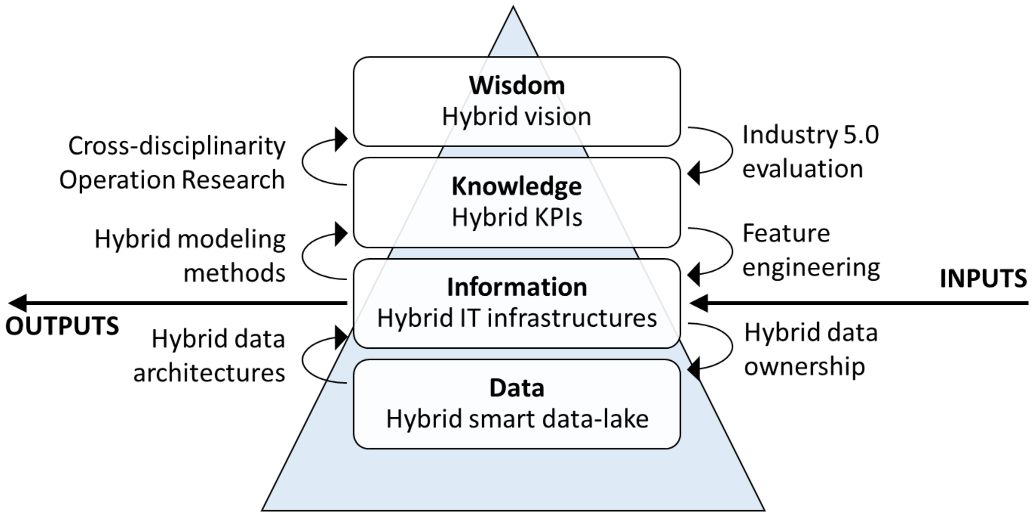

Fig. 3.5 Hybrid-Wisdom structure of the system.

classify different hybridization approaches for different level of IT solutions aimed at extracting value from data: while the scientific problem of DIKW structuring is how to standardize definitions and tools for developing data-driven EIS based on manufacturing knowledge and factory wisdom, the hybrid soul of the framework shifts the scientific point to how selecting and using hybrid modeling techniques in a EIS structured according a DIKW hierarchy. Figures 3.5 and 3.6 briefly resume the framework deeply explain in the following chapters.

This section explain how the framework introduces hybrid techniques into the system through the DIKW modeling used for each agent. Hybrid models, in fact, are presented referred to each level of the DIKW pyramid and change in type based on whether the hybridization is applied to data, information, knowledge, or wisdom. In particular, the hybridization of wisdom is provided by the Manufacturing 5.0 vision, which requires that the goals of an agent operating in a manufacturing system must always consider the centrality of the human being and the sustainability and resilience of the whole system. Such goals therefore underlie the agent's decisions which are the results of hybrid optimization methods, i.e., an optimization problem that uses together all the different kinds of models from operations research that can be considered divided into two main groups: deterministic methods and stochastic methods. Hybrid optimizations choose dynamically at compile time which optimization algorithm to apply from a set of different algorithms that implement the same optimization or find the way to consider at the same time different best solutions

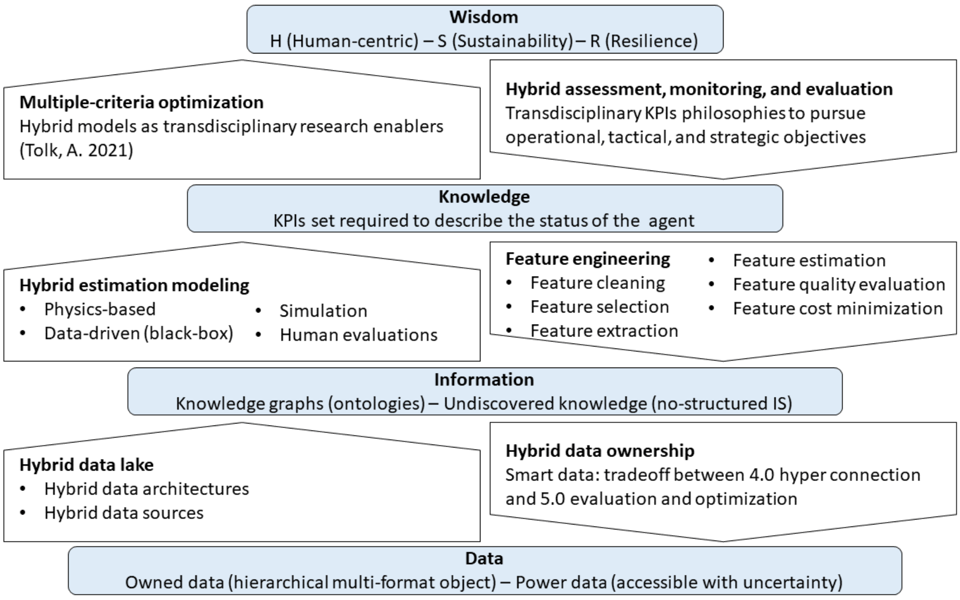

Fig. 3.6 Hybrid level characterization of the DIKW-based system.

from different methods finding the best one. They can use a heuristic to predict the most appropriate algorithm for each piece of code being optimized. [126]

In addition to the use of KPIs referring to different fields, knowledge level hybridization, on the other hand, consists of the models used to estimate the current, and perhaps future, values of these KPIs, i.e., it consists of using inference, physics, expert knowledge, and simulations together to assess the state of an agent at a specific time. This is the main interesting level from the hybridization point of view as it presents the patterns most addressed in the literature. The information level, continuing, lays its hybridization foundations on the fact that hierarchical relational models, structured but more general models represented by ontologies, and unstructured models such as MASs must be used at the same time to describe the relationships between entities in the system. Unlike the other layers, in the information layer it is more difficult to use different models at the same time, but certainly different models are used for different types of agents thus ensuring the hybridization of that layer extended to the entire information of the system.

The following subsections explain in deep the hybridization applied in the information, in the knowledge, and in the wisdom level. For the data level, hybridization consists in the property that each agent has on the data to which it has access: each agent, in fact, owns its own DB (absolute property), while it owns several data "in power", i.e., it has the ability to access in a more or less immediate way the data owned by other agents. These data in potency can be associated with high levels of ownership if the agent communicates directly with the information level of the other agent possessing the data, while they are associated with lower and lower levels of ownership if the number of intermediary agents allowing access to that information increases.

3.3.1 W: cross-disciplinarity awareness

The proposed framework is based on the idea that the environmental cognition and interaction are fundamental, and especially on the hypothesis that the human intervention is (or has to be) the central resource that creates new information structures able of acquiring or generating new knowledge flows. The following ones are consequential requirements for digital manufacturing resources in the industry 5.0 context:

- designed considering all the human resources involved in the manufacturing process,
- focused on the human, or more specific, being a promoter of talent, diversity, and empowerment,
- integrated in the 4.0 sense, i.e. along the functional and the product-service axis interface with the other company's resources and other digital platform internal and external to the company,
- sustainability in the environment in which it operates (the perspective 5.0 is towards a consideration of constraints and needs of the entire planet earth),
- resiliency, or simply, based on flexible and adaptable technologies able to accommodate new information structures and digitize and create new types of awareness.

Considering the manufacturing environment, an industry 4.0 plant usually include various information management systems. The wisdom level, instead, has to be strictly connected to the vision of the manufacturing plant and of all the stakeholders to its activity. However, the wisdom level is designed based on its objective (e.g., the wisdom of an entire MES is so close to the company one, while the wisdom of a simple gateway probably is simpler or does not exist). Figure 3.6 shows the hybrid characterization of the wisdom level with the aim of finding the best trade-off between portability and specific human-centrality, i.e., driven by a specific human knowledge (the systems has to deliver a specific service) and an equal specific human ethics and aims (a specific interpretation of "right" and "wrong").

Human disciplines usually comprise two different focus areas. The first focus looks at the science behind the discipline, dealing with the general principles that build the foundation of the discipline, also known as "the body of knowledge". The second is more interested in finding general methods and solution patterns that can be applied to answer to specific needs of a specific business environmental. They are obviously connected, as methods have to be rooted in general principles to be sure that they will lead to the desired outcome, and new solution patterns may lead to new insights and help to discover new general principles.

Operations Research (OR), as a discipline, has its focus on improvement; hence, it has been argued that the role of OR processes goes beyond the ones explained by a single human discipline: in order to be applied, the OR discipline must surely collaborate with another that is specific of the application field, although several disciplines are probably necessary, for example in manufacturing they can be mechanics, electronics, programming, and marketing. The National Academy of Sciences report on facilitating inter-disciplinary research identified four primary drivers of cross-disciplinarity, [127] namely:

- recognition of the inherent complexity of nature and society, and the inability of reductionism to cope with these challenges,
- exploring problems and questions that are not confined to a single discipline,
- growing societal problems that require a broader approach on a shorter timescale,
- emergence of new technologies that are applicable in more than one discipline,

where simulation is one of these new technologies with the potential to support new forms of collaboration between disciplines.

In the literature there is an interesting work that consider the hybrid modeling approach as an enabler of this trans-disciplinary research, proposing a framework

Cross-disciplinarity as hybrid model of discipline		
Multidisciplinarity	Interdisciplinarity	Transdisciplinarity
Integration		
Parallel, i.e., separated but with the same scope.	In series, or in clear structures.	Synthesis, i.e., complete hybridization.
Communication		
Translation between mapped terms.	Integration of concepts, methodology, procedures, and therms.	Creation of completely new knowledge.
Purpose		
Information exchange: disciplines inform or contextualise each others, often application-oriented.	Knowledge exchange: generating new theoretical, conceptual, and methodological identities, adding cognitive and social aspects, and supporting standardised information exchange.	Shared wisdom: generating common standards and values.

Table 3.1 Key defining features of cross-disciplinary sub-categories. [128]

of Modeling and Simulation (M&S) for the Cross-disciplinarity OR and discussing about the importance of the hybrid modeling for emerging trans-disciplinary areas. [128] Within the M&S community, in particular under the research topic of hybrid approaches, several approaches have been discussed that propose a similar framework to categorise concepts of hybridisation better in support of multi-, inter-, and trans-disciplinary efforts. According this work, "the term "multi-methodology" in OR has been used to describe the combined use of two or more methodologies within a single intervention. It may refers to the combination of qualitative and quantitative methods to more effectively deal with the breadth and nuance of the real world, or to a combination of quantitative methods" of different nature with the aim of combining the benefits or overcoming the weaknesses of individual models.

The human cross-disciplinarity characteristic of the wisdom is considered as the hybrid-wisdom property of decision-making processes of the agent. The terms multi-disciplinarity, inter-disciplinarity and trans-disciplinarity are used to describe different degrees of collaboration of participating disciplines, with multi-disciplinarity and trans-disciplinarity being the two endpoints of this comparison. The term cross-disciplinarity is often used to describe the alignment of vocabularies from different disciplines, creating a common lexicon that can be used in more than one discipline. This framework considers the term **cross-disciplinary** research [128] to mean multi-disciplinarity, inter-disciplinarity, and trans-disciplinarity as shown by the Table 3.1. The cross-disciplinarity guarantees that it is possible to structure the wisdom considering three main characteristics: human-centric, resilience, and sustainability. Sustainability generally refers to environmental, social and economic sustainability. [129]

Transdisciplinary alignment describes the integration of domain knowledge, hypotheses and theories from diverse disciplines. This leads to the development of new composable methods, tools, and applications and new ways of doing research. Transdisciplinary research is challenging for a number of reasons; however, a key aspiration is to share a common language and representation for communication and collaboration. Hybrid models are playing a central role in research that combines the collaboration of more than one discipline. This characteristic of wisdom makes it clear of how the a-priori knowledge of the human is instrumental in defining that level. In this way it becomes clear how, partially answering the third research question (RQ3), knowledge of the human being is implicated in that framework. In addition, as clarified by the discussion of the case study of this work, human agents being very influential in the wisdom obtained as a union of several sources, i.e., such agents being the main sources, it is assumed a strong centrality of the human being in accordance with Vision 5.0 and thus a wide use of its knowledge of the CPS under analysis.

3.3.2 K: hybrid estimations of Key Performance Indicators

The knowledge level can be considered as the level that describes the state of the system, i.e., as the union of all "know-thats" referring to the agent: a DIKW-platform without wisdom still manages to guarantee, for example, a business intelligence service capable of generating value for a company. Such knowledge includes information strictly connected to the agent or referring to external entities and processes in the environment.

During a design phase, the knowledge level for a smart system has to be design considering different types of knowledge. An example is given by the Table 3.2:

knowledge type	**Description**
Domain knowledge	Knowledge for a specified domain. Specialists and experts develop their own domain knowledge and use it for problem solving.
Meta knowledge	Knowledge about knowledge.
Commonsense knowledge	General purpose knowledge expected to be present in every normal human being. Commonsense ideas tend to relate to events within human experience.
Heuristic knowledge	Specific rule-of-thumb or argument derived from experience.
Explicit knowledge	Knowledge that can be easily expressed in words or numbers and shared in the form of data, scientific formulae, product specifications, manuals, and universal principles. It is more formal and systematic.
Tacit knowledge	Knowledge stored in subconscious mind of experts and not easy to document. It is highly personal and hard to formalize, and hence difficult to represent formally in system. Subjective insights, intuitions, emotions, mental models, values and actions are examples of tacit knowledge.

Table 3.2 Knowledge-Based Systems for Development. [130]

The Figure 3.6 shows the characterization of the knowledge. Knowledge within process data, that was largely ignored until the 1980s, has shown a gradual increase afterwards and boomed in the transition to the 21st century, with the emergence of all the conditions that converged to what is now called as Industry 4.0 and Big Data. Models of this kind are called data-driven, statistical, black-box, data analytics, etc., and infer relevant information from large databases. [131] DIKW agent-based model with a balance of smart data and prior knowledge is the base of a general AI service, or better services based on data-driven or machine learning techniques.

The knowledge is a required level to assess, monitor, and evaluate. Such function is the same of a set of KPIs and for this reason the knowledge level can be considered a group of functionalities with the aim of estimating the value of the KPIs of the system considering the values of different informative variables. The variables extracted and selected from the information level are used as input variables to estimate key variables to describe the state of the agent and of other systems in the environment trough an hybrid model using the following individual modelling techniques:

- physics-based or mechanic modeling (classical experimental models),
- data-driven, statistical, or black-box modeling,
- simulation approaches,
- direct evaluation with prior knowledge or done directly by human.

The knowledge level has to deliver the values of the main variables to the wisdom level and such variables can be considered the KPIs set of the agent. This set has to be SMART: Specific ("what exactly needs to be achieved"), Measurable ("how to assess whether the objective is achieved"), Achievable ("the goal has to be attainable"), Relevant ("KPIs have to be significant"), and Time-bound or Timely ("the objective is achieved in a finite time"). [132] In this hybrid level referred to knowledge, the inclusion of direct evaluations by the specialized human as a model to be hybridized with experimental, ML and simulation models, directly answers the third research question (RQ3).

3.3.3 I: information generated by hybrid models

Intelligent systems are not developed separately; rather, they are embedded as modules in a traditional information system to solve tasks related to the intelligent processing of data and knowledge, and this combination, nowadays also often known as a micro-services platform, is referred to as a Hybrid Intelligent Information System (HIIS): an information system based on the idea of "subconsciousness" and "consciousness". [133] The Consciousness Module (CM) is based on conventional data and knowledge processing, which may be based on traditional programming or workflow technology, mainly represented by ontology-based models. They can be classical ontologies, which are developed within the Semantic Web technology (RDF, RDFa, OWL, and OWL2 standards), or nonstandard ontology models including those based on complex networks or object-oriented approach. The Subconsciousness Module (SM) is related to the environment in which a HIIS operates. Because the environment can be represented as a set of continuous signals, the data processing techniques of the MS are mostly based on neural networks, fuzzy logic, and combined neuro-fuzzy methods. The Interaction Module (IM) is added to manage the interaction between the CM and the SM.

The information level is the management level of the interaction with other agents of the DIKW system. So, from the interaction point of view the following options or their combinations are possible:

- interaction is implemented through the SM that processes the data from the environment and transmits them trough the IM to the CM that processes and returns the results that the MS sends to the environment,
- the IM is used for the interaction with another agent and, depending on the tasks to be solved, it can interact with the CM (typically for conventional information systems) or with the SM (typically for systems based on soft computing),
- user interaction can be carried out using the CM (typically for conventional information systems) or through the SM (which can be used, for example, in automated simulators).

Defining this framework, the choice is that both the CM and the SM interact with the environment, but the CM is limited to receive and process only structured information, i.e., information received from well-know agents trough UI, API or directly form the DB of the agent. The SM processes the same data with the addition of data comes from specific API designed to receive unstructured information from the environment to generate new types of knowledge or greater levels of affidability of the knowledge. The IM is used to interact wit the CM in order to send structured information (made such by the other modules) and to modify the CM in order to extend its structure (new types of information are structured and then made processable directly by the CM). The Figure 3.7 shows the structure of the information level composed by these three main modules.

Resuming, the CM performs (i) the data processing for knowledge generation on the ground of ontology-based models, (ii) the logical control and consistency check of the data which are received from the IM,and (iii) the implementation of new components in order to receive of new types of inputs and output. The SM performs (i) the data mining for knowledge generation on the ground of ontology-based models and (ii) the receiving and processing of not-strictly-structured information. The IM performs (i) the management of the use of the SM to generate structured information to send to the CM and the implementation of new components of the CM that needs to process new types of information.

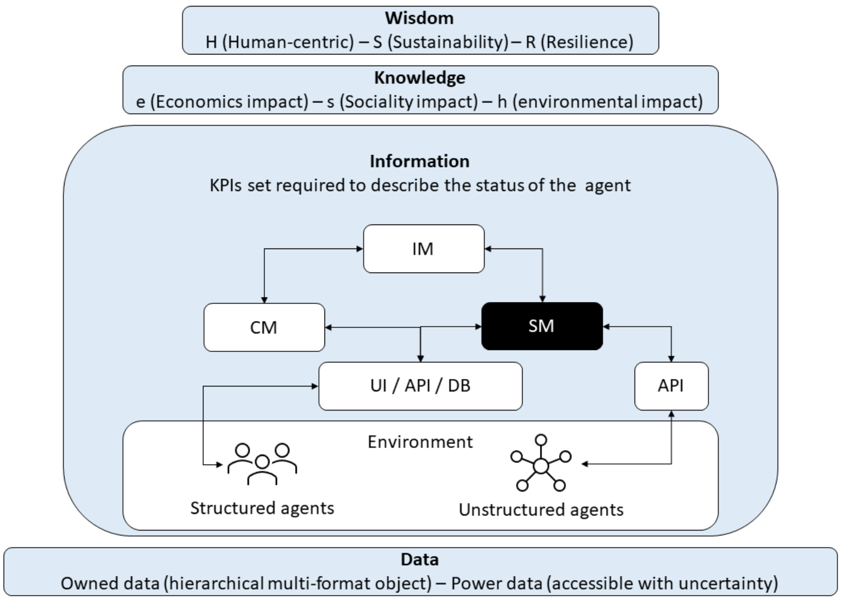

Fig. 3.7 Scheme of the information level with the Consciousness Module (CM), the Subconsciousness Module (SM)and the Interaction Module (IM).

3.4 Consideration on the HW-framework

3.4.1 KPIs for the framework

A Key Performance Indicator (KPI) is a metric used to evaluate whether or not a system is meeting its objectives. By definition, not all metrics can be "key", so KPIs are a select grouping of metrics deemed essential to meeting business objectives. Within IT, KPIs are very effective for answering the following questions: are we making investments in the right places, are we getting the results we expect to see, and are our plant, especially human resources, satisfied with the value they have received?

In the case of IT system design assistance frameworks such as the case of the framework in this thesis, or in general in the case of KPIs for creative activity, it is often difficult to measure performance and in fact it is rare to find work of this type in the scientific literature. The development of metrics, indeed, is very difficult in areas of high abstraction such as software development and the general design process. [134] The following points indicate why better estimation of effort and duration of design projects is so important:

- in some situations, cost or schedule overruns lead to project termination,
- schedule overrun increases the risk of product obsolescence due to the increased risk of missing the market window (in many cases this means a project failure),
- an initial delay in a project can engender further delays. [135]

From the literature, the main factors to consider designing a CPS are manpower, [136] complexity, amount of change, time trough each design phase, development cost, clarity of contents, [137] design difficulty and resources. [138] Following these works and adding other typologies of metrics, the KPIs selected for this framework are the following ones and divided in project characteristics, outcome, and development process metrics:

- **designers number**, i.e., the number of designers employed in the design phase,

- **complexity** in therm of functional complexity and technical difficulty (an example of metrics used to calculate this KPI is the number of agent of each type),
- **amount of change**, i.e., the number of modifications to implement in the environment of the systems (interfaces of other systems, work habits, and more),
- **time** trough each design phase (introduction, formalization time, transmission of concepts to developers, total time),
- **completeness**, i.e., whether the framework can model all the resources needed for CPS or not,
- **scalability**, i.e., how easy it is to generate new agents (even of different types) so as to increase the CPS in therm of entities and resources,
- **flexibility**, i.e., the ease with which such framework can be transported between different application areas with different characteristics,

The metrics described are generic KPIs, and the part about the primitive metrics from which these KPIs are calculated was not addressed. This part has not been addressed because it is strictly dependent on the application of the framework, i.e., the area of application of the CPS being designed. These indicators were constructed considering to evaluate the performance of the use of that framework in a particular case study, but they can be used, considering different case studies, to evaluate the performance of the design framework in general in order to compare it with other techniques in the literature.

3.4.2 4.0 technological context

This framework is intended to be a design tool in the 4.0 manufacturing context, and its use is possible because technologies and methodologies specific to this era are employed in the development of the designed CPS. In particular, the digital twin concept presented with a DIKW characterization can be considered an optimum formalization for an agent referred to a product. In fact, the digital twin is a technology that by definition has several hybrid models at its core in order to obtain accurate estimates regarding the current and future states (or referring to hypothetical scenarios) of products and manufacturing resources.

Another technology that enables such frameworks is the IIoT. This technology, now increasingly widely used, has spread design thinking based on decentralized or agent-based systems, that is, systems composed of increasingly micro services as opposed to the monolithic systems of the previous era. In the 4.0 context, therefore, software designers have a greater capacity for abstraction needed to design MAS. To confirm how the IoT philosophy can support the application of such a framework, consider that an IoT architecture is composed of five main layers, which can be called the (i) physical layer, the (ii) sensor and actor layer, the (iii) connectivity layer, the (iv) analytics layer and the (v) digital service one. [139] These five layers can be easily compared to the layers of the DIKW structure associating the sensor and the actor with data, the connectivity with information management, the analytics with knowledge, and the service with the wisdom of the agent.

Finally, data analysis and simulation methods, in the technological context of Big Data, make it possible to assume a constant availability of sufficient data and models to estimate all the variables needed to describe the entire state of an agent, that is, to estimate all the KPIs present in its knowledge layer and their projections in future or hypothetical scenarios. Data analytics helps manufacturing firms to get actionable insights resulting in smarter decisions and better business outcomes [140]. For this reason, data analytics is becoming a very attractive topic for almost every manufacturing firm in Industry 4.0 era. Generally the data analytics is covered under three sub topics. First one is descriptive analytics that summarizes the data and reports the past. It answers the question "what has happened?" and extracts information from raw data. [141] There is also an extension to the descriptive analytics named "diagnostic analytics" which reports the past but tries to answer the questions like "why did it happen?". [142] Second sub topic is predictive analytics which is considered as the forecasting phase. It answers the questions "what will happen?" and "why will it happen?" in the future. [141] These two sub topics are exactly the function of the knowledge level that describes the states of the referring agents. Finally, goal of the prescriptive analytics is to provide business value through better strategic and operational decisions. It is all about providing advice. In general, prescriptive analytics is also a predictive analytics which prescribes some courses of actions and shows the likely outcome or influence of each action. It answers the questions "what should I do?" and "why should I do it?", [142] and this is strictly connected to the knowledge of the consequences of the decision made considering the actual wisdom of the agent.

In addition to being enabled by the technologies of the 4.0 era, this framework is designed to provide support for the design of a CPS, that is, to support the design of a digital system in accordance with the paradigms of the fourth digital revolution (as well as the principles of Industry 5.0 as explained in the following section). For disciplines such as cyber security, it is convenient to have a detailed description of the entire telecommunications system obtained through the description of the information level of each individual agent. In this method, security experts can obtain the information about the sensitivity of the data by considering the associated metadata or by considering the frequency of use of that data by all agents in the system. In addition, thanks to the agent-based philosophy, in the information layer, it is possible to describe type-specific security protocols for agents based on the sensitivity of the information they contain. With regard to IIoT and cloud computing, the micro-service structure of this framework makes the application of such technologies immediate. Vertical and horizontal integration is an inherent paradigm in the conceptual foundations of the model, and should it be used to model the entire reference CPS it would ensure such integration. If, on the other hand, it is used for sub-components of the entire system, however, it will allow easy integration via the API described in the agent information layer or via specific agents created precisely with the aggregation API functionality for different agents in the system. Finally, as far as collaborative robots are concerned, the wisdom layer is useful in that it can be extended to them as well and can be used in order to provide the vision that aids robot decision-making systems with a clearer and more comprehensive description of business goals.

3.4.3 5.0 view context

This paragraph is the last one that provides final considerations on the framework and, in particular, contextualizes this work at the dawn of manufacturing 5.0. As explained in the introductory paragraphs, this progress revolves around the concepts of human-centric, sustainability and resilience, so it is analyzed below how this framework supports this vision based on these three main points. Before analyzing these points one by one, it is important to underline how one of the main reasons behind this framework is to allow the introduction of a vision common to all agents that can act as a guide for all the different types of decisions that this system will have to support. This concept is the same foundation that led researchers and sector experts to spread the term Industry 5.0. In fact, this fifth revolution does not substantially involve any new technology or methodology, but underlines how the use of the most advanced technologies risks being ineffective or, worse, unsustainable if there is no application criterion that considers the interests of the entire ecosystem. Indeed, innovation ecosystems need to be governed, and cannot be left alone to their own course. [143] [87] The decisions concerning the selection of conceptual frameworks that inform innovation ecosystem governance are important because they influence what, why, where, how, and for whom the innovations materialize or not. Industry 5.0 is about building complex and hyper-connected digital networks without compromising long-term safety and sustainability of an innovation ecosystem and its constituents. The level of wisdom is therefore the characteristic of this framework that allows software designers to introduce this vision into the systems they design.

Considering the first research question (RQ1), the three pillars of the Industry 5.0 are considered has three main KPI groups to evaluate the activity of the agent: how it is human-centric, if it is sustainable for the system and how it contributes to the whole sustainability of the system, and how it contributes to resilience. In addition to this, the centrality of the human being lies in the prior knowledge which represents one of the main methods of generating knowledge on the state of the agent to be integrated with data-driven and physics-based (and simulation) methods through hybridization. Sustainability, on the other hand, is addressed through the principles of smart data and ease of integration with the other EISs present in the company. Retroactively, in fact, the level of wisdom evaluates the use of only the data strictly necessary for the system, finding the right trade-off between managing the large amount of data provided by 4.0 technologies and the use of small data which, however, often do not guarantee the levels of accuracy necessary for the production of knowledge able to optimize manufacturing processes. As regards the ease of integration, this framework gives the necessary freedom to design a digital system customized to the needs of the factory in such a way that it can be both a single platform (composed of various micro-services) that carries out all the functions necessary for management, and a platform that can be integrate with the EISs already present in the company (such as ERP systems, MESs, WH management systems and PLM platforms) in order to provide the functions that these cannot perform. Finally, the resilience of the system is provided through the agents' ability to generate new types of knowledge, or at least to underline the need, through the structure of the information level presented in the previous paragraphs. Furthermore, the entire system of agents is not only scalable (easy addition of agents of the same type as other existing ones), but it can be extended to new types of agents without modifying the existing structure in such a way as to satisfy the need to insert new manufacturing resources, new production processes or new methodologies and paradigms in factory activity.

Chapter 4

HW-TPM system for CNC tool machiness

The aim of this case study is to propose a hybrid system that takes advantage of the strengths of both methods (physics-based and data-driven) while minimizing the effect of their weaknesses. The main maintenance terms and concepts used in this chapter are referring to a Total Productive Maintenance (TPM) system according lean thinking and 4.0 paradigms. Summarising, the case study is focused on the design of a Hybrid Wisdom-based TPM (HW-TPM) system. The proposed framework is applicable to the maintenance of a generic tool set on a generic CNC machine, like welding machines, milling machines, 3D printers and other machines based on the same technology. In other works, this chapter considers manufacturing processes executed by machines with cyclical jobs, i.e., which can be monitored in a fixed defined time window.

After the introduction section, the rest of the chapter is organized as follows: the second section defines the proposed hybrid model for a TCM system of CNC tool machine, the third is the case study based on open data for wear predictions of milling tools, and, finally, the last section presents the conclusions and the ideas for future improvements.

4.1 Introduction

4.1.1 Maintenance processes

Maintenance costs are estimated as a percentage of production costs that vary between 15%, for the manufacturing sector in general, and up to 40% for the metal-working industry. [144] With the proper implementation of Predictive Maintenance (PdM) strategies, these costs can be reduced by up to 30% [145], by automatizing part of monitoring activities, and by optimizing the decision of replacing the resources when strictly necessary. Furthermore, a PdM strategy can reduce the incidence of failures by up to 70%, allowing the productive time of systems to be increased by up to 30%. [146] Another important estimation is that PdM methods based on Machine Learning (ML) algorithms can reduce current maintenance costs by an additional 30%, increasing machine operating life and reducing downtime. [147]

In the era of Industry 4.0, several technologies, as Internet of Things (IoT) and Artificial Intelligence (AI), enable the real-time data collection required by high-performed PdM methods. In fact, they are based on real-time data collected by sensors systems and Big Data infrastructures. [148] Research has proposed several AI-based methods whose performance grows as the information possessed about the process under observation increases, but they are inoperable with no real-time in-formation. The performance of physics-based methods depends on (i) the number of variables (sources of variability) considered in the model, (ii) the complexity of the physical laws, and (iii) the estimation quality of few parameters with data offline generated by experiments and inspections. Contrarily, the accuracy of data-driven methods depends on the quantity and quality of historical data, which are difficult to replicate for research analyses. [149]

Starting in 1950, preventive maintenance was introduced in order to limit the effects of a failure, which with the previous approach often led to downtime of the entire production process. [150] Differently, it has been estimated that 99% of mechanical failures can be predicted with the help of specific indicators, on this basis was born the Condition Based Maintenance (CBM) [151] to overcame limitations given by a simple strategy based on time or age schedules. It involves two main processes: diagnostic and prognostic. [152] The improvement of these methods is represented by PdM models, in which measurements on the tool machine are used in combination with process performance data measured by other devices. The use of such data jointly allows statistical models to analyze historical trends in order to predict the instant when the machine needs an intervention. [153]

From the simple visual inspection of the machine, the failures prevention has evolved to automated methods that use traditional signal processing techniques and new Machine Learning (ML) methods. The maintenance strategy optimizes a trade-off situation between maximizing the useful life of a component and up-times through early replacement of this component (time-based PM), which has been demonstrated to be ineffective for most equipment components, considered as flawed and unreliable in recent years. [144] A good estimation of the tool status avoids the use of degraded tools that reduce the work surface quality and excessive preventive replacements, which involve higher costs and production time. The paradigm of Industry 4.0 proposes the digitalization and the interconnection of machines, thus improving the possibilities of having a more effective condition monitoring, also by analysing the data coming from sensors.

Prognostics methods can be categorized into data-driven, physics-based and hybrid approaches [154]. Despite the significant recent progress in the Model-Based (MB) and ML hybrid modeling domain, there are various challenges that throttle down the full-fledged growth of hybrid modeling: (i) there is no guidelines for selecting hybrid models, (ii) there are few benchmarks (problems and dataset) for evaluating and comparing hybrid models, (iii) training accurate models with low amount of data or labels, (iv) minimizing data collection costs, (v) solving the complexity due to geometric data formats as CAD files and imbalanced data. [1]

4.1.2 Case study justification

The aim of this work is to propose a model of CPSs designed to monitor the status of the tool set on a generic CNC machine. Even if the term “Tool condition monitoring” refers to the monitoring of different conditions of the tool (e.g., wear, breakage, did-not-cut condition, thread depth, damaged or missing threads), we limit our analysis to the prediction of wear and Remaining Useful Life. The proposed framework can be applied to a generic machine tool for two main reasons: (i) all types of wear of the tool consist in a loss of material that is measurable with standard methods, and (ii) each wear type produces increasing of forces and temperatures to which the machine answers by varying the work status, measurable trough a standard set of sensors.

Data-driven PM has been extensively applied to industrial manufacturing using machine learning algorithms, such as logistic regression (LR), support vector machine (SVM), decision tree (DT) or random forest (RM), and neural networks (NN). [155] However, previous works mainly addressed single type of sensor measurements, they are focused on one specific learning algorithm, and the majority of them used data from private experiments or generated from a simulator. Furthermore, few applications addressed the maintenance of cutting machines, since most works are related to fault detection of bearings or motors.

Differently from the previous works, the case study is focused on more scientific problems defining a formal framework for data-driven PM that can be applied to every CNC machine to support the tool replacement planning. [156] To this aim, a great attention was firstly devoted to present a formal definition of variables and parameters applicable in the domain. Data cleaning and feature manipulation procedures are central parts of this work and they are distributed among agents. Other agents have knowledge levels based on hybrid techniques of data selection, outlier detection, feature extraction, feature normalization and feature selection that result a novelty with respect to the state of the art. [2]

The case study is focused on the development of a Tool Condition Monitoring (TCM) service whose main component is the knowledge level describing the wear condition of milling tools provided by physics-based (white-) and data-driven (or black-box or ML) models based on real time sensor data. The aim is the development of a Total Productive Maintenance (TPM) CPS designed according the proposed framework. The milling process is a scenario assumed in order to have quantitative results for models used in the "hybrid-Knowledge based sub-system". A milling process case is used, at the expense of others such as welding, mainly because it is based on open data, provided online by NASA.

4.2 HW-TPM general framework

Designing a CPS that performs a Condition-Based (CB) Predictive Maintenance of a single manufacturing CNC tool machine. The proposed framework has the aim of supporting the design of a general platform managing digital agents referable to physical manufacturing resource (in particular, a smart digital twin of a machine tool) that is able to assess, monitor, and evaluate (or contribute to the evaluation) its status, its performances, its impact on the system and preferable decisions to make by and regarding the agent towards a TPM system.

The first step is to define the overall agent, the entire system, so as to obtain the definition of the wisdom level that is the presentation of the entire project. The following subsection are structured in order to be a sort of template for a project design document of an hybrid platform to support maintenance choices enabled by different types of model with the same main aim of create a sort of digital twin of the tool, or in other words with the same aim of use historical information to estimate the actual and future states of the machine tool. [4]

4.2.1 General description of the HW-TPM system

The Figure 4.1 tries to resume the main characteristics of the system.

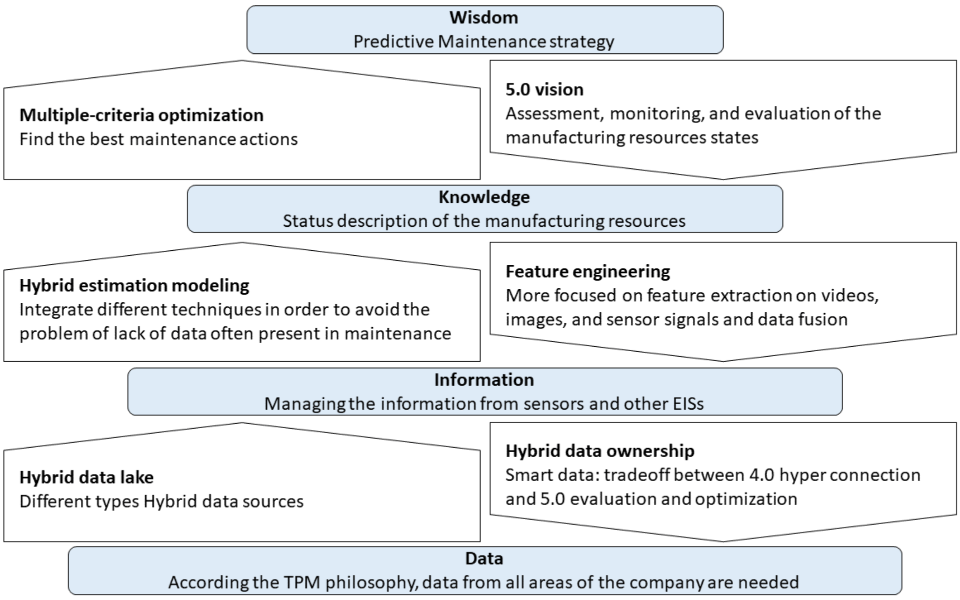

Fig. 4.1 Hybrid level characterization of the HW-TPM system.

Project description. The aim of the project is to develop a Decision Support System (DSS) for Condition-Based (CB) Predictive Maintenance of CNC machine tools, i.e., a platform able to manage data from MES and maintenance inspection office, in order to support the decision of substituting or not the tool of a single autonomous CNC machine used for the surface finish of rings produced by an additive machine.

Goals of the project The project is referring to a monolithic platform providing a specific family of services:

- acquiring real time sensors measurements from the 4.0 CNC tool machine,
- acquiring production parameters values and tools characterization from the MES of the plant trough specific APIs,
- acquiring wear measurements provided by maintenance inspections and communicating with them trough the plant chatting service,
- estimating the actual wear status of machine tools pursuing the goal of no needed direct measurements,
- informing the machine operator regarding the risk level of working with the set tool,
- delivering forecasted future wear states,
- estimating the Remaining Useful Life (RUL) or lifespan of the tool set in the machine.

Project justification. Maximize the utilization of the tool life, and avoid using worn tools that are considered detrimental to the quality of the part. In addition, the system to be developed contributes to optimal inventory management, work planning and scheduling based on resource status, and product quality certification.

Agents. The agents of the system, and an idea of their relationships, are shown by Figure 4.2 and they are the following ones:

1. **tool**, that represents the main resource for this system, i.e., the one to monitor and to manage respect its wear, and it is able to (i) evaluate the best time to be substituted, (ii) ask for a further inspection by the metrology expert, and (iii) update the tool's stock availability,
2. **machine**, that is the manufacturing machine on which the tool is set,

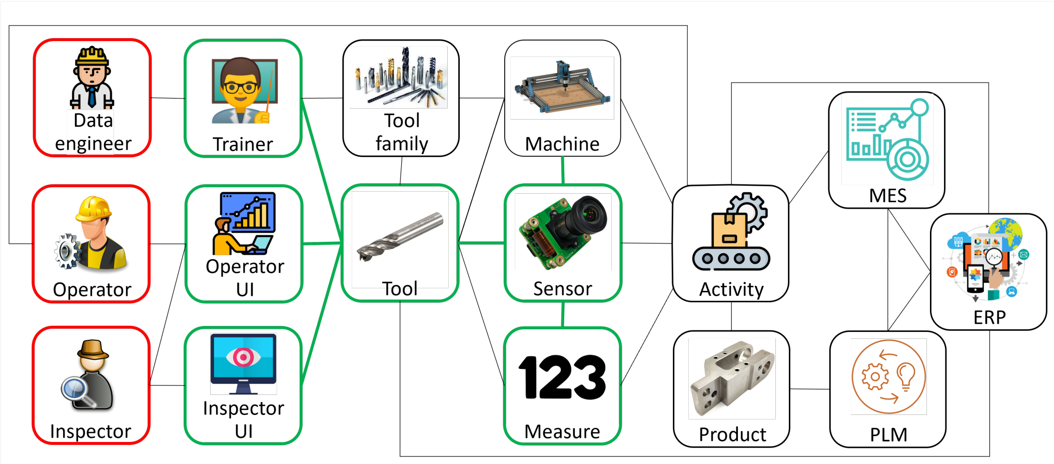

Fig. 4.2 Agents involved in the HW-TPM system. The main agents of the system are underlined in green,while agents referring to HRs in red.

3. **sensor**, that is a component of the machine (in case of 4.0 machine or an external sensors used during the machine activity) and for example can be a thermometer, an accelerometer, a thermal camera, a video camera, an AC or DC sensor, a microphone for decibels measuring, or any other type of sensor for monitoring the activity of the machine and its components,
4. **measure**, that is a single measurement generated by a single sensor during a specific activity of the machine, i.e., digital signals (time series), pictures, videos, row data and all possible data generated by sensors,
5. **trainer**, that is the agent responsible for controlling the process of estimating the tool wear level and its Remaining Usefull Life (RUL),
6. **tool family**, that is he family to which the tool belongs and which describes its general characteristics,
7. **activity**, that is the operations run by the operator on the machine to produce the desired product (component),
8. **MES**, that is an EIS used in the plant and it control each manufacturing activity,
9. **ERP**, that is an EIS used in the plant,
10. **product**, that is the component produced with the machine activity,
11. **PLM**, that is an EIS used in the plant to monitor the entire product life cycle,

12. **operator**, that is the human resource needed to start a task on the machine and to replace the tool as needed,
13. **operator UI**, that is the UI component used by the machine operator,
14. **inspector**, that is the human resource responsible for directly measuring the level of tool wear,
15. **inspector UI**, that is the UI component used by the inspector (metrology expert),
16. **data engineer**, that can interact with the trainer agent directly modifying its code source.

4.2.2 Agents description: measure

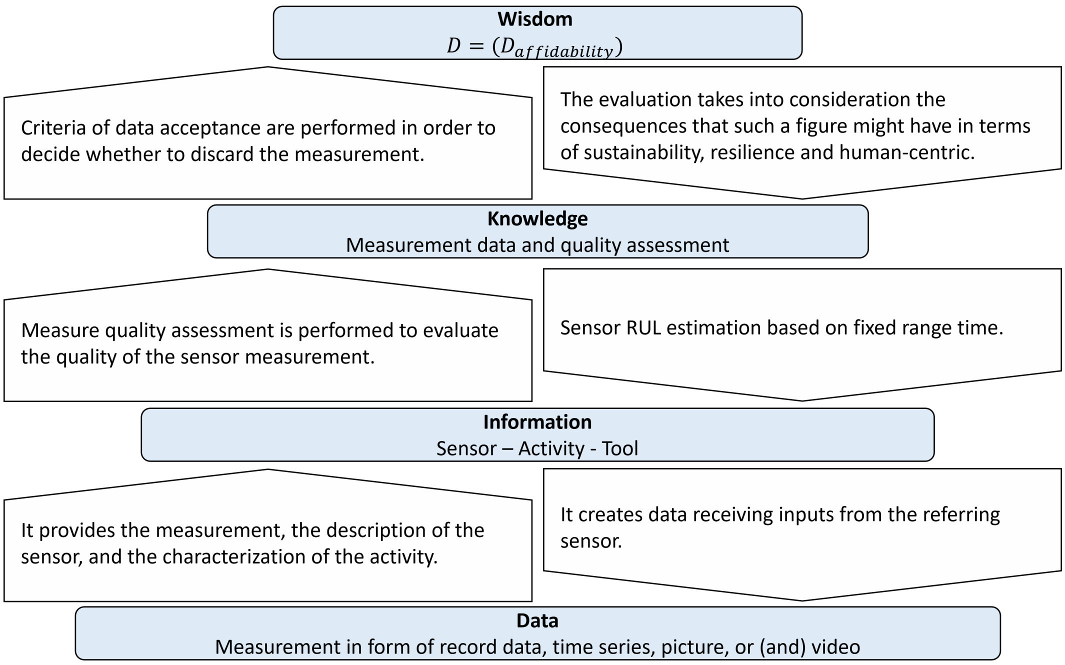

Fig. 4.3 Agent characterization for the measure.

The measure, shown by the Figure 4.3, is a complete agent consisting in a single measurement run by the corresponding sensor and it is link to the tool agent in order to estimate its wear level and to the activity agent in order to deliver summarising statistics describing the performed activity. The level of knowledge of such an agent

Ref.	Measurement	Model	Machine	Data source
[157]	Acoustic emission, vibration	SVM	Bearings	Private experiment
[158]	Vibration	LR, SVM	Bearings	Simulation and public dataset
[159]	Vibration	ANN	Bearings	Private experiment
[160]	Vibration	LR, SVN, ANN	Bearings	Private experiment
[161]	Temperature, pressure, speed	SVM	Aircraft engine	Private experiment
[162]	Vibration	SVM	Bearings	Private experiment
[163]	Vibration	SVM	Gearbox	Private experiment
[164]	Electric	SVM	Vehicle subsystem	Private experiment
[165]	Vibration	LR	Bearings	Private experiment
[166]	Vibration	LR	Punch	Private experiment
[167]	Acoustic emission	LR	Cutting tool	Private experiment
[168]	Optical	ANN	Laser welding	Private experiment
[169]	Vibration	DNN	Bearings	Private experiment
[170]	Pressure, lubricant	DNN	Gearbox	Private experiment
[171]	Vibration	DT	Spur gear	Simulation
[172]	Acoustic emission, vibration, force	DT	Cutting tool	Private experiment
[173]	Acoustic emission, vibration	Improved LR	Cutting tool	Public dataset
[174]	Temperature	DT	Refrigerant flow system	Private experiment
[175]	Vibration	DT	Gearbox	Private experiment

Table 4.1 Research applications in the literature with data characterization: sensors, model, Machine and data source. [128]

consists of one or more metrics of the quality of the data, i.e., of the entire sensor measurement, so as to provide all the resources necessary for the level of wisdom to estimate what is the best choice between holding or not holding the measurement (eliminating or not itself as a corrupt agent).

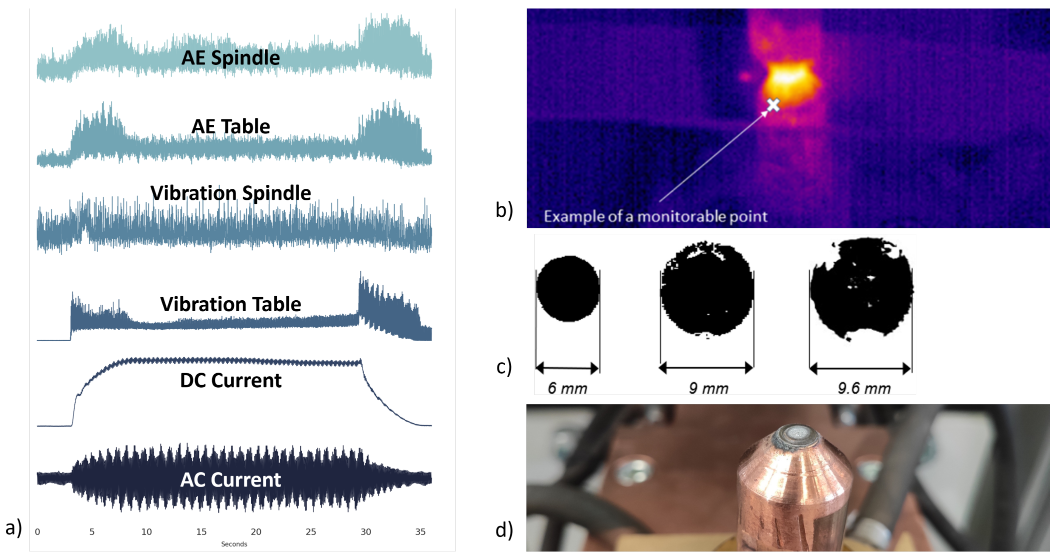

Fig. 4.4 Examples of data formats and types referring to numeric time series measurements, images and videos, and high-cost manual measurements: a) milling data set, [176] b) thermal video of electrode, carbon imprint image, [5] and a tool picture. [177]

Data level. Table 4.1 and Figure 4.4 show examples of type of data that are outputs of sensors used to measure a certain physical amount. Such data may be:

1. single data (e.g., liters of lubricant used) for which no processing is necessary, just domain checks;
2. simple record data or one-dimensional data, like numerical time series representing the measurement of a certain quantity during the machine's activity with a certain frequency (e.g., vibration of the work table, alternating current, the noise generated by component processing, or or other sensors shown by the Table 4.2), for which it is necessary to apply dimensionality reduction techniques,
3. images or videos recorded during operations (e.g., videos produced by a thermal video camera or photographs of the most critical areas of the tool that attempt to capture wear represented in the form of a spot or otherwise irregular areas on the surface), for which classical and non-computer vision techniques are performed in order to extract image-related statistics such as average or maximum temperature, or number of worn areas or their average diameter;
4. results of tests performed in the laboratory or otherwise using expensive procedures that block the availability of the tool (e.g., tribological measurements),
5. complete unstructured information (e.g., text messages or comments written in the operator UI), for which specific text mining or Natural Language Processing (NLP) are required.

Information level. Data structure consists in a time series structure in case of numerical list measurements, images and videos formats, and customized other structure for general high-cost manual measurements (often performed by human). The information level link the measurement to a single sensor, tool and activity. Trough this level, the sensor agent could report the rejection of the measurement.

Knowledge level. The processing of measurements is composed by 3 steps:

- acquisition, trough sensor API (metadata included in sensor agent),
- domain data validation, i.e. the analysis of each sensor output to identify over domain outliers and wrong values, consisting in checking that values are in their sensor domain,

Equipment	Vibration	Humidity	Ambient temperature	Ambient pressure	Acoustic signal	Thermography	Motor current	Insulation resistance	Electrical capacitance	Electrical inductance
Pump	✓		✓	✓	✓	✓	✓	✓		
Valve		✓		✓	✓					
Motor/Fan	✓		✓		✓	✓	✓	✓		✓
Heat exchangers	✓	✓	✓	✓						
Steam turbine	✓	✓	✓	✓	✓					
Electrical and electronic equipment			✓			✓		✓	✓	✓
Cables and connectors			✓			✓		✓	✓	✓
Pump seal		✓		✓	✓			✓		
Piping/Structures	✓				✓					
Compressor	✓				✓	✓	✓			

Table 4.2 Sensors used to predict failures and wear of different manufacturing resources according the literature. [178]

- time series outlier detection should be a no-invasive detection method able to detect only extreme measurements considered a symptom of emergency situation for humans or other resources.

Wisdom level. The wisdom level is strongly related to the sensor agent and its wisdom level and it is consequently focused on the concept of affidability and on the problem of reject or not the measure, that is reasonably based on the percentage of missing values, but should also be function of the following indicators toward Industry 5.0: human measurement efforts and risks, level of adaptation of control parameters to different tool families (resilience score), and sustainability related to the rejection or validation of the measurement (whole agent).

4.2.3 Agents description: sensor

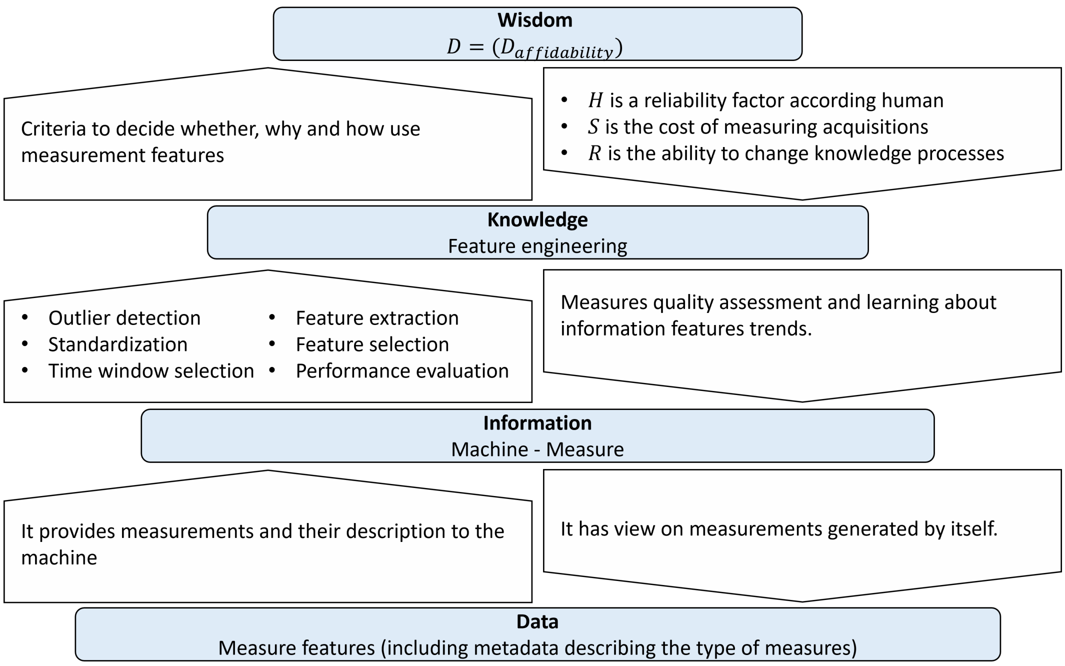

Fig. 4.5 Agent characterization for the sensor.

The information level of a sensor agent, Figure 4.5, ties each measurement to the corresponding machinery in which it is mounted. The sensors performs a feature engineering process that make available high level information about measurements. In the proposed framework, the distinction between searching for outliers within a sensor measurement and between measurements is considered, to recognize extreme signals that indicate process instability. [2] [3] [4]

Data level. This agent contains metadata needed to understand how read the corresponding measurements, variables extracted from sensors measurements and feature performance statistics.

Information level. Sensor measurements are linked to a certain activity and to a certain machine. The communication with measures, machine, activity and tool agents are managed by API on the same server. But data are received from the API powered by the CNC machine seller. Before sending data to the knowledge level, as shown by Figure 4.6, (i) intra-measure outlier detection, (ii) missing measures management and (iii) time window selection activities are performed.

1. The intra-measure outlier detection removes single points from the set that composes a single sensor measurement,

2. while the missing values is assumed in this work as a simple check of missing values came from data acquisition or outlier detection activities.

3. Performance measures, like percentage of outliers or managed missing values, are created as meta-data of the output: in this work, missing values do not affect the knowledge level activity and for this reason they are ignored. In case the knowledge level rejects the measurement according the missing values intra-measure percentage, the information level transmit the information to the measurement agent.

4. The stationary time window selection is the selection of a sub-window of the time series that is strictly connected to the machine activity. For this purpose, a Change Point Detection (CPD) technique is suggested, i.e. a technique to identify instants in which the probability distribution of a time series changes. [179] [180] In each measurement there different phases of the machine: the (i) available phase in which the machine is on but not working, (ii) the start and the end of processing and the (iii) stationary phase that is in the middle of them. To do this a multiple CPD is run on each signal, obtaining multiple sub.series. The stationary sub-series is the one with maximum value for the average, in case of non-negative asymmetric signals (e.g., the direct current), or for the standard deviation, in case of symmetric signals (e.g., the alternating current).

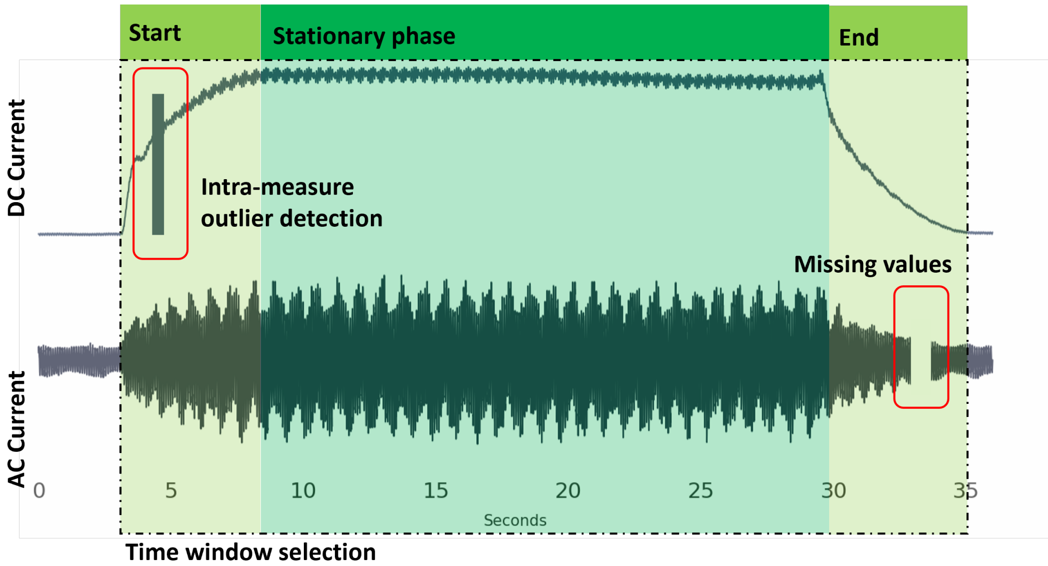

Fig. 4.6 Functions performed by the information level of the tool agent.

Knowledge level. In this level, (i) measurement acceptance, (ii) data normalization, (ii) feature extraction, (iii) inter-measurements outlier detection, (iv) feature selection, and (v) performance evaluation activities are performed. This process consists in the hybridization of the data sources considering the overall best features: the hybrid system consists in the data fusion process. [181]

1. A measurement is accepted and then processed if the percentage of missing values is below a certain threshold, or rejected and reported to the information level in case. The case of rejected measurement (missing complete time series) is estimated and reported to the information level. The missing sensor outputs causes a missing feature set and this statistics set are estimated according two cases: (i) if its the first or the second tool activity, they are the mean statistics of the first or second activities performed by tools of the same family; (ii) from the third activity onward they are predicts considering the previous statistics referring to the same tool: time series approach or simple linear regression are suggested.

2. Data normalization is used to improve the overall quality of a data avoiding situations in which some values over-weighting others. A general choice for the normalization is a Min-Max normalization. [182]

3. The feature extraction function is need for the first main dimensionality-reduction contribution to the problem: from an order-of-magnitude number of points certainly in the thousands or more, to a limited number of statistics describing the measurement. Single missing measurement points do not affect this function because it was sufficient not considering them for statistics calculation.

4. The inter-measurements outlier detection is performed comparing a set of statistics of the single measurement with the set of cumulative statistics of other measurements of the same sensor. The standard deviations comparison is a first suggestion to understand if a measurement is an outlier agent or not. Another method could be the classical outlier detection based on the Inter-quartile Rule for Outliers [183] where the Inter-quartile Range (IQR), calculated on a specific selection of other measurement agents, is multiplied for $k_IQR = 1.5$ and an agent results an outlier according this method.

5. The goal of the feature selection is to reduce the dimension of the feature set, in order to both save training time and minimize the over-fitting risk. [184] The feature selection presented in this work is completely unsupervised (i.e. executed without considering target variables) in order to guarantee the independence from tool agent information. This function is activated periodically, usually based on the number of new data (measurements) received, and keeps the agent updated on which statistics are most meaningful for the purpose of describing a measurement.

When performing analysis of complex data, one of the major problems is dealing with the high number of data involved. The purpose of feature extraction is to reduce the initial set of measured data, by extracting only the essential and explanatory features, thus simplifying the subsequent learning phases. [185] The quality and quantity of features are key determinants which highly affect the result of the prediction. The feature to be extracted belong to the following three types: time domain, [65] frequency domain [65] and polynomial regression coefficients.

- *Time domain statistics.* Feature extraction in time domain allows to evaluate the magnitude of the signal. It consists in collecting a set of statistics that describes the stationary time series. Examples of time domain feature are maximum value, mean value, root mean square, standard deviation, Skewness, Kurtosis, peak-to-peak and crest factor. Time domain features are important, but they only reflect the signal changes over time. For this reason, frequency domain features need to be extracted too.

- *Frequency domain statistics.* The frequency-domain [186] is the power spectrum of the stationary time series and it describes the distribution of power into frequency components composing. The signal digitized as a time series can be converted in the frequency domain by using the Discrete Fourier Transformation (DFT) applicable to finite sequences of equally spaced samples and statistics are calculated on module and argument of the complex outputs of the DFT. Another list of statistics can be defined for the band power spectrum: examples of statistics are maximum, sum, mean, standard deviation, Skewness, Kurtosis, and relative spectral peal. [65]

- *Time polynomial regression coefficients.* Finally, the last feature extraction technique consists in a forward selection procedure to select a n-grade polynomial regression of the time domain series with a fixed maximum degree n_{max}, considering not only the stationary phase but all the time series. The coefficients of this regression polynomial are considered as a subset of features.

Wisdom level. The wisdom level is strongly related to the sensor agent and its wisdom level and it is consequently focused on the concept of affidability and on the problem of reject or not the measure, that is reasonably based on the percentage of missing values, but should also be function of the following indicators toward Industry 5.0: human measurement efforts and risks, level of adaptation of control parameters to different tool families (resilience score), and sustainability related to the rejection or validation of the measurement (whole agent). Having multiple feature available, based on the evaluations made in this level, the knowledge level perform feature selection processes towards wisdom targets. A resilience strategy could be to periodically test even excluded features (because until today not relevant) or new features (for example adding new types of sensors) to check their performance: this is a re-training phase managed by the trainer agent.

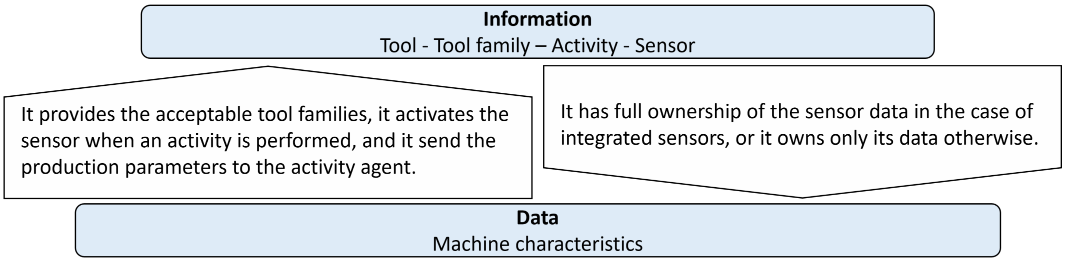

Fig. 4.7 Agent characterization for the machine.

4.2.4 Agents description: machine

The machine and sensor agents change their relationship based on whether or not the sensor is integrated into the machinery. In a 4.0 manufacturing scenario is can be assumed that sensors are physically mounted in the machinery, therefore, the sensor agent has a strong affiliation with the machine system.

The Figure 4.7 shows the structure of the CNC **machine agent**. In this case it is simply an information system with the automatic scope of delivering information regarding which types of tools can be set on it (selection of tool families) and it activates the sensor agent when an operation is performed. Finally, when a task is performed the machine agent send to the activity one the information about the production parameters used to process the product and any error, warning or simple description message about the activity just performed.

4.2.5 Agents description: tool

As shown by the Figure 4.8, the tool is the most complex, or smartest agent. It communicates with the machine to receive all information related to the manufacturing activity (start, end, type of activity, type of product if necessary, and possibly all information coming from the MES or PLM system), with the sensors to receive data from the sensors related to its activity and extract relevant features (and with the measures for singular cases), with the trainer to be supported in constantly updating predictive models, with the two agents referring to the graphical interfaces with operator and inspector, and, finally, with the tool family to keep all the characteristics regarding its mechanical and economical properties and awareness of inventories related to other equivalent tools. [2] [3] [4]

Data level. Data on process parameters, sensors, and wear levels are used to estimate the current wear levels of the specific tool and its RUL (lifespan), i.e., an estimate of how much tool life the tool has so as to optimize tool management choices and similar resource stocks. These estimates are made by hybridizing three types of models:

1. first, any reports are handled by the operator on the machine whose experience is a valuable company asset,
2. secondly, a model based on experimental physical relationships that, given the same process parameters and tool characteristics, describe the trends in different measures of tool wear,
3. and, finally, a data-driven model based on ML methodologies that can relate tool wear to measurements obtained from sensors in such a way as to obtain more accurate real-time estimates because they are based on the peculiarities of the individual tool's history.

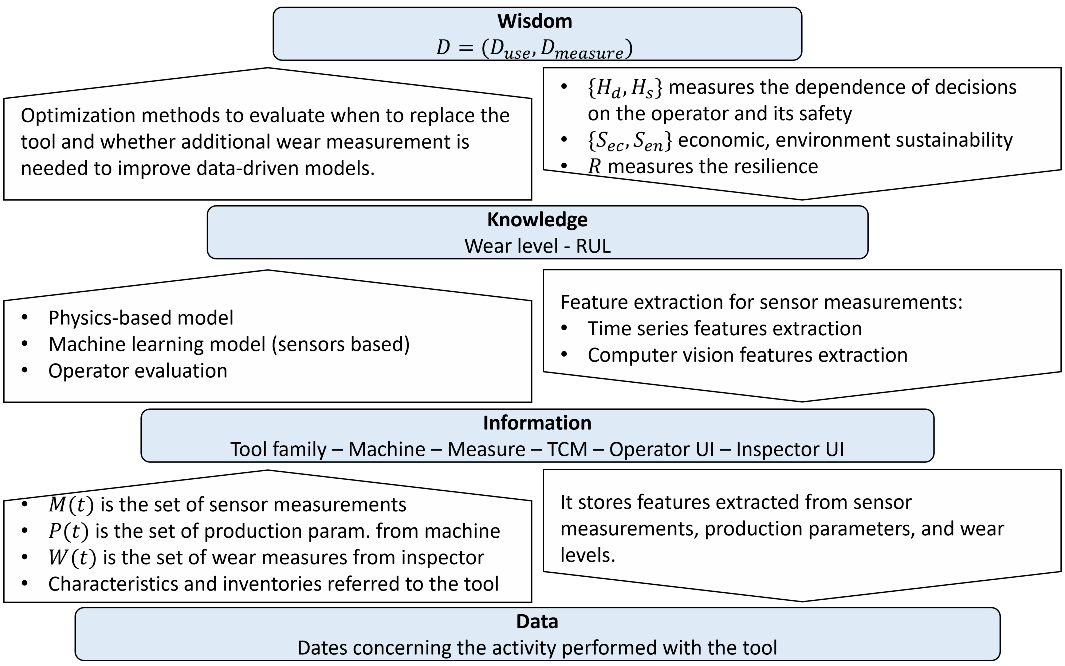

Fig. 4.8 Agent characterization for the tool.

Information level. The observed manufacturing process regarding the tool is modeled as a multivariate discrete time stochastic process that takes into account different target variables describing the wear level of the tool or if the tool is usable or not. [187] Generally, wear mainly depends on (i) tool and work piece material, (ii) tool geometry, (iii) manufacturing parameters, (iv) type and level of lubrication and (v) machine-tool characteristics. The monitoring of tool wear is important as the wear affects decreased accuracy of produced parts, poor surface finish, economics of cutting operations and decreased tool life. Tool wear is an important factor to estimate since it is not directly measurable without stopping the process and perform an inspection of the work piece. [188] Having to enable interaction with the trainer agent and the knowledge layer with ML algorithms, this layer continues the logic of managing the division of data into training and test sets, and, if necessary, validation sets.

Knowledge level. The ML model, therefore, is the only one capable of incorporating the information contained in sensor measurements within the features it uses to estimate target variables. The problem type is a supervised learning with wear direct measurements as target variables and production parameters, tool and product characteristics and sensors measurements as sources to obtain the predictor variables. Mechanistic or phenomenological knowledge is the more sophisticated type of explicit knowledge as it involves deep understanding of established laws of physics, phenomena explained by physical or chemical causes or relations between empirical observations of phenomena. It is often described by mathematical models. These are called first-principles, mechanistic, phenomenological or white- box models. Generally used interchangeably, the aforementioned terms actually refer to sightly different modeling approaches: (i) first-principles models come entirely from the established laws of nature; (ii) mechanistic models describe systems by referencing the laws of nature, but their parameters, although with physical meaning, must be determined from data; (iii) phenomenological models represent empirical relations and their parameters determined from data do not have physical meaning; (iv) finally, white-box models refer to situations where all relevant interactions are known and are made explicit, in a transparent way. [189] For the sake of simplicity, the term "physics-based model" is assumed to address to any of the citations described above. The proposed method in the knowledge level is a hybrid model between physics-based and data-driven approaches, and it aims to exploit the potential of each method. [4] To these two models, evaluations (no measurements) given by the operator and its experience, i.e., its knowledge about the processes, are added as third estimation of the wear.

To understand how to hybridize such models, an example referring to the union of models based on physics and ML techniques is presented below. The first step consists in training the physics-based model and data-driven (black-box or machine learning) model individually, i.e., a parallel hybrid structure is used. In this way, the two methods generate estimations about wear levels for each activity executed with the same tool, called $W_{PB}(t)$ and $W_{ML}(t)$, respectively. Then, always with the training set, for each run, the optimal weight $\omega(t)$ is calculated to generate the linear combinations of physics-based and data-driven predictions as stated in the following equation: $W = \omega W_{PB} + (1-\omega)W_{ML}$. The weights are defined by choosing as objective to minimize the Root-Mean Square Error (RMSE) and the Root Relative Squared Error (RRSE) of the hybrid predictions. Finally, the trained hybrid model is evaluated on the test set. The estimation of the wear level W during the operation performed at time t with the same tool is a weighted average of physics-based and data driven methods. However, the model estimates a dynamic weight with which to manage the use of different wear prediction models in order to be able to "trust" more of the best ones over the life cycle of the tool, i.e., to choose which model to give more weight based on the number and characteristics of operations performed by the tool. This example of hybrid modeling, one of the simplest ones, can still be extended for a general hybrid system with several wear estimation models: an example is the hybrid model $W = \omega_{PB}W_{PB} + \omega_{sens}W_{sens} + \omega_{sim}W_{sim} + \omega_H W_H + \omega_{NLP}W_{NLP} + \omega_M W_M$, where W_{PB} is the estimation given by a (or more ensemble) Physics-Based model, W_{sens} by supervised ML methods applied on sensor data , W_{sim} by simulation methods, W_H by estimation given by manufacturing humans, W_{NLP} by ML methods applied on free comment given by humans, and, finally, W_M are time series or, in general, auto-regressive methods applied on direct measurements of target variables, i.e., wear measurements.

As from literature, the RUL is estimated on the actual wear level of the tool compared to a fix threshold limit [3] or to past decisions taken by machine operators or technical specialized operator, or based on experimental data or other information resources, in other words, the RUL is estimated considering the actual knowledge state regarding the tool and the past decisions according wisdom level criteria. Considering that each tool performs the same type of task for its entire life cycle by changing only materials and production parameters (in other words, always running the same milling G-code), an estimation of RUL is the number of jobs still performable before the tool, that carried out some jobs, does not respect quality parameters. [65] The Useful Life is a stochastic variable referring to the number of jobs for a new tool before exceeding quality limits: from the $(UL+1)$-th activity is the case of a bad maintenance strategy (the tool is used even if its UL is finished, thus causing a decrease of product quality), if the UL-th activity is decided to be the last one the case refers to an optimal maintenance strategy (the tool is changed exactly when needed), and, finally, if the tool do not perform all the activity until the UL-th one, this is the case of a too preventive maintenance strategy (the tool is changed when it could have worked for one or more jobs). Concluding, a suggested model for RUL is the one that, for the first manufacturing activity, assigns the UL value calculated as average of the tool family, and for the following activities a linear model is used to forecast the future wear level in order to allow the tool agent to calculate how many runs (activity) are estimated before reaching unacceptable wear levels. The parameters of the linear model are calculated using a training dataset which will include both the wear values that are certain, i.e., measured by the inspector, and the wear values predicted by the tool agent itself.

Wisdom level. Estimates regarding wear and RUL are then used by the tool to assess whether or not it needs to be replaced through optimization methods that can be simply represented by control thresholds (if wear is likely to exceed a threshold value then replace) to more complex methods that can consider production scheduling, inventory inventories, and all costs, direct and otherwise, related to tool activity and any misuse of the tool. Such tool management support is provided to the operator at the machine via related UI. A further decision managed by the tool is to request an intervention by the employee to directly measure the wear level in order to provide an additional target data that can improve the supervised ML algorithm. This decision is made on the basis of the interaction with the operator who provides discordant estimates, or considering that the tool is replaced in very different times than those suggested, or finally because, interacting with the trainer agent, the tool perceives that its forecast performance estimates are lower especially when compared to tools belonging to other families. The decisions made by the tool are then evaluated from a vision 5.0 perspective, i.e., from a human-centrality perspective, where $H = \{H_1, \cdot, H_n\}$ represents different metrics related to that pillar, from a sustainability perspective, $\{S_i\}_{i=1,\cdot,n}$, and from a resilience perspective, $\{R_i\}_{i=1,\cdot,n}$. In the Figure 4.8, there are two examples of metrics for human-centrality and sustainability. Having multiple models available, based on the evaluations made in this level, the knowledge level decides which ones to implement or, rather, which ones to give weights greater than zero. A resilience strategy could be to periodically test even the models with zero weight so as to recheck their performance: this is a re-training phase managed by the trainer agent.

4.2.6 Agents description: trainer

The trainer agent is strictly connected to the tool one. ML models, in fact, need to be trained to be operational and to re-run this training process every time you have new data and want to update model parameters. Fully automated supervised training without human intervention is a complex function that would therefore overload the tool agent and for this reason this function is relegated to the trainer agent. That agent therefore, as shown by Figure 4.9, checks the tool's predicted performance by comparing it with pre-set benchmark performance and with the performance of other tools from the same and different families, and evaluates whether to perform the training phase to update the model parameters by managing the use of all useful data in the system, i.e., by choosing from the data on all tools and their performance. The trainer agent is, therefore, responsible for updating the hybrid sub-systems, or levels, that make up the DIKW structure. The updating of the wisdom level models and structures is not addressed in this paper.

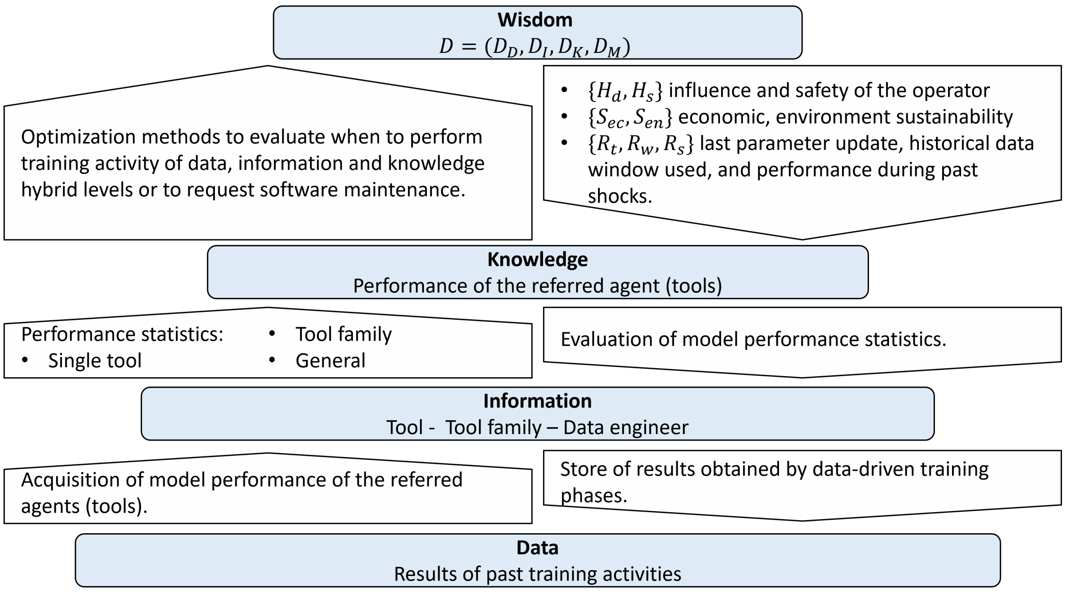

Fig. 4.9 Agent characterization for the trainer.

Data level. Results and performance statistics of past training activities, that are strictly link to the training phase and not to the performance of such trained models used by the tool agent. While the results are the performance obtained on the test set, examples of training performance statistics are: training time, over-fitting risk rate or the energy consumed during such phase.

Information level level. The trainer communicates with the tool family and the tool agent in order to upload their information level, that for mostly consists of resuming the use of some abandoned features or inserting new ones to re-execute the feature selection phase as a data fusion method. Updating the information system means updating information sources always tends to reward a hybrid solution in order to maximize the current quality of knowledge generation processes. This agent communicates directly with the data engineer

Knowledge level. This trainer level is the most complex, as it has several complex training subsystems. In fact, training such a hybrid system means both (i) training the hybrid model and the way how merge different predictive models and (ii) training each individual model. For example, physics-based models may need to upload the values of some constants; for models based on sensor data, the classic ML supervised training phase is required; simulations are retrained more rarely for singular scenarios, a request for estimation from the operator is a training task, training of NLP methods follows the theories of deep learning, and finally, the agent trainer may require a direct measurement of wear and tear if it estimates performance such that the number of certain data needs to be increased in order to improve the predictive power of the supervised algorithm. When training different models, a model validation procedure is used to minimize the probability of over-fitting. An example of a model validation technique is K-fold cross-validation, which involves splitting the training set into K smaller sets of the same size (fold), in order to train the model k times on the training set given by the union of the remaining of $K-1$ folds and validate it on the set given by the exclude fold. The final measure of model performance is then the average of the K performances obtained.

Wisdom level. Based on the evaluation under the 5.0 view of the performance of the agents under control (tool) and their training processes, this level manages the training of each hybrid level of the DIKW structure of the dependent agent and also provides, in case where adequate performance cannot be achieved, a direct request of intervention by a data management expert in order to modify the code source of the agent. One proposal for indicators from the perspective of 5.0 is to consider the rate of influence of human-input information, the rate of safety of operators in working with current management performance, the cost and impact that each training phase has, and, for resilience, all those factors that describe how up-to-date a model is based on new available data, which time window of data is considered to assess its performance and what that performance is, especially in case of shocks, i.e., how long (or how much data) the models took before they reached the performance they had before a significant change in the system that caused lower estimation performance.

4.3 Case study on milling open data

The steps described in the previous section were applied to the use case dataset. All the analyses were conducted with the free software environment R (R Foundation) and the aim is to predict the flank wear coefficient, shown by Figure 4.10, using data generated by sensor monitoring the machine milling activity. Flank wear occurs at the tool flanks, where it contacts with the finished surface, as a result of abrasion and adhesion wear. The cutting force increases with flank wear. It affects the great extent of mechanics of cutting. The flank wear region is known as wear land and is measured by the width of wear land. If the width of wear land exceeds $0.5 - 0.6mm$ the excessive cutting forces cause tool failure.

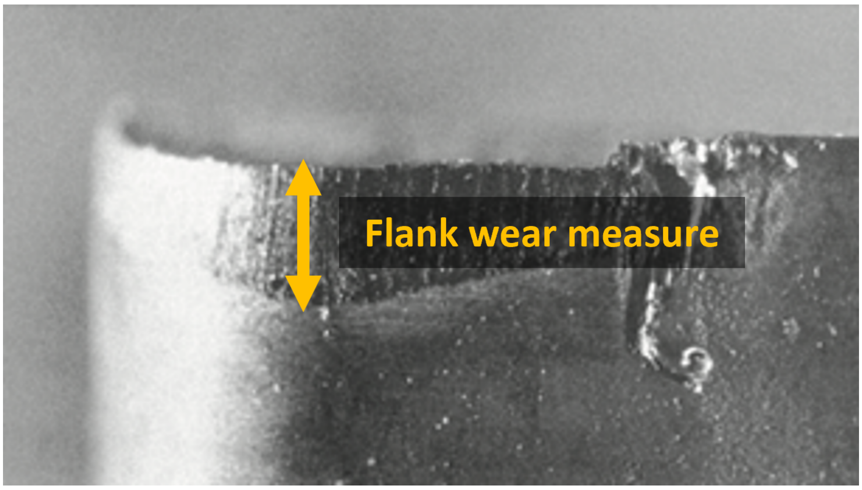

Fig. 4.10 Flank wear caused by friction between the flank face of the tool and the machined work-piece surface and leads to loss of the cutting edge.

4.3.1 Experimental system: NASA milling open dataset

Unfortunately the data level is not considerable hybrid, the information one is limited by the laboratory environment (few typologies of measurement), and the wisdom level consists in a quick analysis of the UX/UI of the machine tool operator with the maintenance functionalities of the wisdom system and the integration of the system itself with a more general enterprise information system, like a PLM system, an ERP one or a MES.

The analyses performed in this case study are based on the public Milling dataset, made available by the Prognostic Center of Excellence NASA-PCoE [176]. It contains the values recorded by 6 sensors throughout the life cycle of 16 tools (cases), under different working conditions identified with 8 scenarios, for a total of 170 machining operations (runs). Each case is characterized by 3 machining parameters, which follow the recommendations of the tool manufacturer, and by the type of material machined with fixed dimensions ($483mm \times 178mm \times 51mm$). Except for the cutting speed, which remains unchanged at $200m/min$, the other variables are dichotomous and in particular: the feed rate has been fixed at $0.25mm/s$ or $0.5mm/s$, the depth of cut has been fixed at $0.75mm$ or $1.5mm$, while the work-piece materials are cast iron or stainless steel 145.

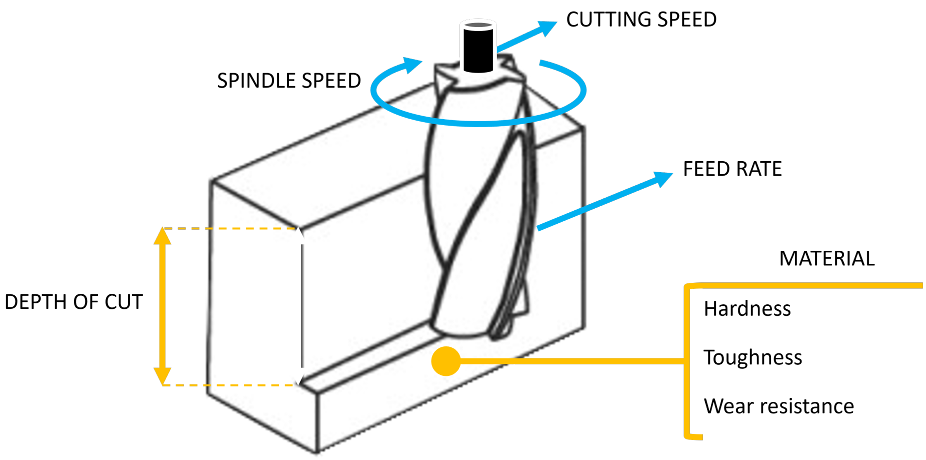

Fig. 4.11 Process parameters representation.

The design of the experiment involves three production parameters shown by Figure 4.11: (i) workpiece material, (ii) feed rate and (iii) depth of cut. Particularly, the investigated materials were (i.a) cast iron and (i.b) stainless steel 145, the feed was set to (ii.a) 0.25 mm/s or (ii.b) 0.5 mm/s and the depth of cut (DOC) was set to (iii.a) 0.75 mm or (iii.b) 1.5 mm. The choice of parameters was guided by industrial applicability and recommended manufacturer's settings. The Flank wear was computed as the distance from the cutting edge to the end of the abrasive wear on the flank face of the tool. After each run of the experiment, the insert was taken out of the tool and the wear was measured by a microscope. Flank wear is the most commonly observed and unavoidable phenomenon in metal cutting, which is also a major source of economic loss resulting due to material loss and machine down time, thus it is a generally accepted parameter for evaluating tool wear. [190]

For each parameters combination there are results regarding 2 experiments so as to obtain 16 cases analyzed. For each case, there are some number of milling activities (runs) executed with the tool. In each run of each case, 6 sensor measurements were collected: the alternate and the direct motor currents ("smcAC and smcDC"), the vibrations and the acoustic emissions of the table and of the spindle ("vib_table", "vib_spindle", "AE_table" and "AE_spindle"). The data window of each sensor is a sample of 9000 time-ordered elements. Since the total number of runs is 170, the total number of signals recorded is 1020. Table 4.3 shows the complete Design Of Experiment (DOE) used to generate this open dataset.

The agents that interact in that CPS are:

1. **inspector**, that, with a special microscope, measures the flank wear coefficient removing the insert from the cutter and measuring the distance from the cutting edge to the end of the abrasive wear on the side face of the tool., that is the human resource responsible for directly measuring the level of tool wear,
2. **inspector UI**, that is the csv file representing the DOE,
3. **tool**, that is changed by the inspector,
4. **sensor**, tha are
5. **measure**, that is a single measurement generated by a single sensor during a specific activity of the machine, i.e., digital signals (time series), pictures, videos, row data and all possible data generated by sensors,
6. **trainer**, that is the agent responsible for controlling the process of estimating the tool wear level and its Remaining Usefull Life (RUL),

Case	N. runs	Material	Feed	DOC
1	17	Cast iron	0.5	1.5
2	14	Cast iron	0.5	0.75
3	16	Cast iron	0.25	0.75
4	7	Cast iron	0.25	1.5
5	6	Stainless steel	0.5	1.5
6	1	Stainless steel	0.25	1.5
7	8	Stainless steel	0.25	0.75
8	6	Stainless steel	0.5	0.75
9	9	Cast iron	0.5	1.5
10	10	Cast iron	0.25	1.5
11	23	Cast iron	0.25	0.75
12	15	Cast iron	0.5	0.75
13	15	Stainless steel	0.25	0.75
14	10	Stainless steel	0.5	0.75
15	7	Stainless steel	0.25	1.5
16	6	Stainless steel	0.5	1.5

Table 4.3 Values of production parameters for the 16 milling processes of the use case. Note that the production parameters using in the first half experiments are the same of the second one.

7. **tool family**, that is he family to which the tool belongs and which describes its general characteristics,
8. **activity**, that is the operations run by the operator on the machine to produce the desired product (component),
9. **MES**, that is an EIS used in the plant and it control each manufacturing activity,
10. **ERP**, that is an EIS used in the plant,
11. **product**, that is the component produced with the machine activity,
12. **PLM**, that is an EIS used in the plant to monitor the entire product life cycle,
13. **operator**, that is the human resource needed to start a task on the machine and to replace the tool as needed,
14. **operator UI**, that is the UI component used by the machine operator.

4.3.2 Measure agent

The agents relating to time series measurements are assumed, without loss of generality, having the same sampling size and sampling frequency, i.e. the distance between two consecutive sensor acquisitions. This means that milling activities are assumed of the same type, or at least with a similar nominal duration, which allow to use a single width for all the measurement windows with only few cases with truncated acquisitions managed by the outlier detection phase of the sensor agent.

All the measures in this system are assumed lists of 9000 thousands numerical values. For the data validation, we used the domain $(0,+\infty)$ for all the sensors except the alternate motor current for which it was set to $(-\infty,+\infty)$, that are all the possible values that physical quantity can theoretically reach. We checked the time series data and we found that no one of them exceeded the fixed thresholds, thus all data are acceptable. All measurements of the dataset are accepted from this agent.

4.3.3 Sensor agent

The outlier detection identified 33 signals as outliers, over the total of 1020 signals recorded. The 33 outliers are reported in Figure 4.12 on 6 rows, one for each type of sensor. For example, for the alternate motor current, three signals were detected as outliers, the 18-th one (corresponding to the first run of the 2 second case), the 95-th one (corresponding to the first run of the eleventh case), and the 115-th one (corresponding to the 21 run of the eleventh case). For the direct motor current, two signals were detected as outliers, again the 18-th and the 95-th. As it is clear from the figure, it appears that if the signal collected in a specific run by a sensor is an outlier, also the signals collected by the other sensors at the same time may be outliers. After the application of the outlier detection procedure, only the four signals related the 95-th measurements were removed. For the other measurements, the result of the outlier detection is shown always in Figure 4.12

For the stationary window selection, the cpm-package in R language [191] is used. For the identification of the first change point a the function "detectChangePointBatch" is used with a confidence level $\alpha = 0.1\%$. The method used for the change point detection are the Bartlett test statistic for the alternating current and the

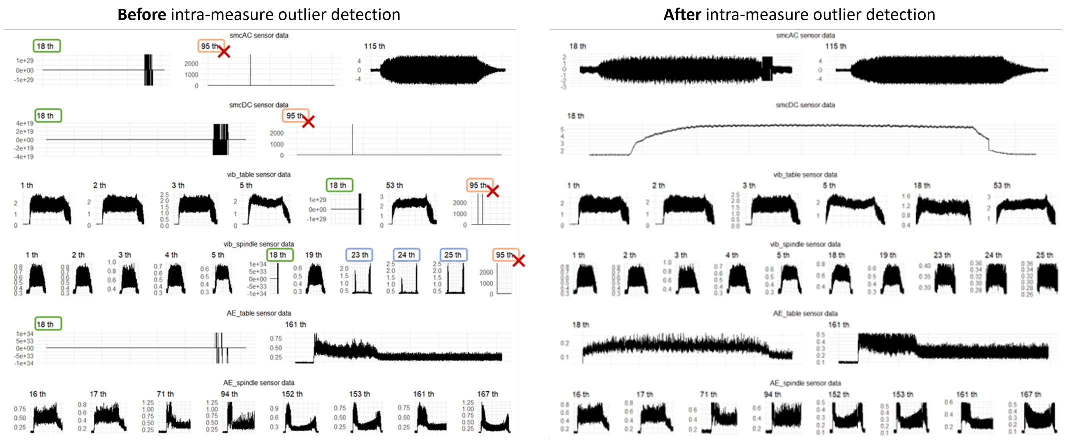

Fig. 4.12 Plots of sensor measurements identified as outlier to be processed and sensor measurements identified as outlier but accepted (after outlier detection).

Generalized Likelihood Ratio test for the other sensors. In Figure 4.13, an example of the data change points identification on the six sensors measurements is shown.

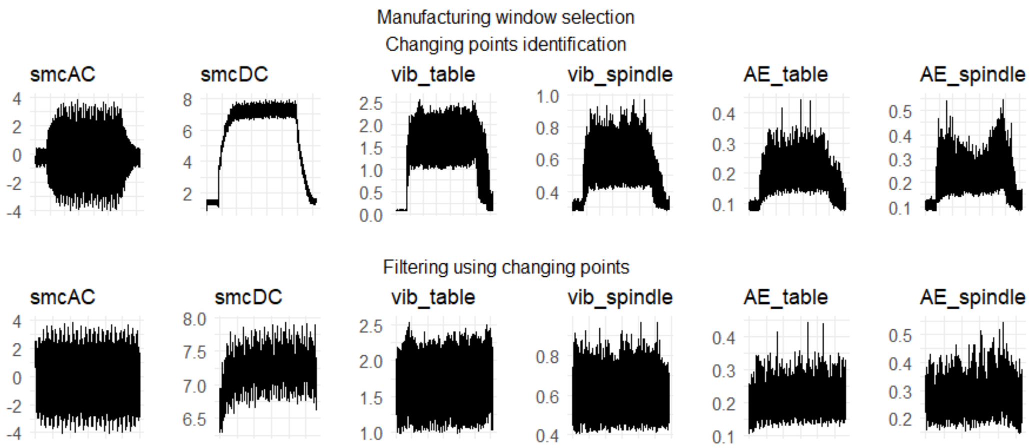

Fig. 4.13 Stationary window selection applied on Milling Dataset.

The three types of features (i.e., time domain statistics, frequency domain statistics and polynomial regression coefficients) were extracted from each sensor measurement. The statistics over time domain are the following: (1) maximum, (2) minimum, (3) mean, (4) standard deviation, (5) root mean square, (6) skewness, (7) Kurtosis and the (8) crest factor. The statistics over frequency domain are calculated on the module and the argument of the complex outputs of the "fftw" package [192] and they are the following: (1) maximum, (2) minimum, (3) mean, (4) standard deviation, (5) Skewness, (6) Kurtosis and the (7) Relative Spectral Peak per Band.

Finally, the coefficient of regression polynomial with maximum degree of 5 are computed, starting from the constant polynomial. In each i-th iteration the new polynomial is considered, until finding the first i_{max}-th iteration in which the p-value referred to fitting a polynomial of maximum degree equal to i_{max} is significant with confidence level of 5%. Summarizing, the number of features extracted from each measurement is 21, with respect to the dimension of a measurement equal to 9000: 8 features from time-domain, 7 from frequency domain and 6 as coefficients of polynomials of degree 5, where for polynomials of lower degree the last coefficients are equal to zero.

For the normalization, the Min-Max strategy was used, e.g., a linear transformation to map each feature to the set $[0,1]$.

The proposed feature selection methods are: (1) removing features with a low coefficient of variation, (2) selection by correlation between features, (3) selection by hypothesis testing, (4) selection by monotonicity and prognosability. Other methods could be the selection by trendability or multicollinearity analysis. In the first step, for each feature the coefficient of variation is calculated and, if it is under the threshold value of 0.5 (it is recommended a value less than 0.75 [193]), the feature is no longer considered. This is to remove constant or low variability features. In the second step, the correlation between each couple of features is calculated and if the absolute value of this correlation is greater than 0.8 (an arbitrary value of positive correlation) the feature with minimum coefficient of variation is no longer considered. Finally, an ideal prognostic feature has two quality: monotonicity and prognosability. [194] These parameters are calculated for each feature and a weighted average of them is compared to a minimum value: features with average lower than this value are removed.

4.3.4 Tool agent

As the tool proceeds in the job execution, the various deterioration mechanisms, such as abrasion and plastic deformation, result in increasing levels of wear on the tool surface, and as consequence, in the reduction of its RUL. The gradual wear can be divided into two basic types, corresponding to two regions in the cutting tool. [195] Flank wear occurs on the relief face of the tool and it is measured by the width of the wear band VB. Crater wear consists of a cavity in the rake face of the tool that forms and grows from the action of the chip sliding against the surface and it can be measured either by its depth or its area. All the measures of wear are called direct measurements in [196].

ML methods can utilize the information contained in sensor measurements can be used to estimate tool wear. In the proposed case study is analyzed in the literature. [4] [3] [2] To estimate the tool wear, ML models have 126 features: for each 6 sensors are used the 21 feature extracted from a single measurement. A total of 22 missing values was detected. They were estimated as explained in the previous section. For the normalization, the Min-Max strategy was used. ML models require supervised [197] using a training set and a test set. The ML algorithms implemented in the case study are the most used in literature [198] and the only one to be employed, i.e., which output is considered as ML wear estimation W_{ML}, is select as the best one between the following:

1. Linear Regression (LR),
2. Bayesian Linear Regression (BLR) where linear regression in which the data are supplemented with additional information in the form of a prior probability distribution,
3. Decision Forest (DF) that is an ensemble model that operates by constructing a multitude of regression decision trees at training time,
4. Boosted Decision Tree (BDT) where boosting means that each tree is dependent on prior trees,
5. Neural Network (NN).

In parallel, classical mechanics methods concerning the wear of rotating machine tools are based on the non-linear relationship between two main parameters: cutting speed V_c and T_{RUL} (RUL of the tool considering the amount of past manufacturing activities). A first equation was proposed by Taylor in 1906 and has the following form: $V_c \cdot t^{1/\beta} = C$ and $C = \alpha f^{-C/\beta} d^{-\gamma/\beta}$, where C is the cutting speed with which one minute of life is obtained, d is the depth of cut, f is the feed rate and α, β and γ are empirical constants. $1/\beta$ is an indicator of how much the tool life is affected by changes in cutting speed and empirical data has defined that $1/\beta \in [0.1; 0.15]$ for HS steel tools, $1/\beta \in [0.2; 0.25]$ for carbide tools and $1/\beta \in [0.6; 1]$ for ceramic tools. [199] The physics-based component of the hybrid model (called W_{PB}) is based on the extended Taylor equation for rotary tools, which includes all machining parameters. The equation has the following form: $T_{UL} = \alpha_0 v_c^{\alpha_1} f^{\alpha_2} d^{\alpha_3} W_{PB}^{\alpha_4}$, where $[T_{UL}] = [min]$ is the estimation of the useful life of the tool where $T_{UL} = T_{RUL}(t = 0)$, $[v_c] = [mm/min]$ is the cutting speed, $[f] = [mm/rev]$ is the feed rate, $[d] = [mm]$ is the dept of cut. $[W_{PB}] = [mm]$ is the width of flank wear according to the physics-based method that can be measured in relation to the activity time T, and α_0, α_1, α_2, α_3, and α_4 are empirical constants to be estimated that will be called coefficients, since they represent the coefficients in the multiple regression model. The formula can be rewritten to obtain the estimation of wear level like shown by the Equation 4.1.

$$W_{PB}(T) = e^{\ln(T)-\ln(\alpha_0)-\alpha_1 \ln(v_c)-\alpha_2 \ln(f)-\alpha_3 \ln(d)} \tag{4.1}$$

The estimation of the coefficients can be done with a multiple linear regression by performing a logarithmic transformation, which can be written in matrix form like in Equation 4.2.

$$Y = \begin{bmatrix} \ln(T_1) \\ \vdots \\ \ln(T_n) \end{bmatrix} = X\alpha = \begin{bmatrix} 1 & \ln(v_{c,1}) & \ln(f_1) & \ln(d_1) & \ln(W_{PB,1}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln(v_{c,n}) & \ln(f_n) & \ln(d_n) & \ln(W_{PB,n}) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}, \tag{4.2}$$

where n is the number of operations rune with the same tool, X is called sensitivity matrix, [200] and α contains the coefficients that can be estimated with the method of least squares. The accuracy of the estimation of coefficients depends on the inverse of the matrix $X^T X$, the term on which the optimization criteria are based, as for example to define the minimum data sample to be collected for training the method. [201]

Then, always with interacting with the trainer agent, for each run, the optimal weight $\omega(t)$ is calculated to generate the linear combinations of physics-based and data-driven predictions as stated in the following equation: $W = \omega W_{PB} + (1 - \omega) W_{ML}$.

4.3.5 Trainer agent

To test the stability of the algorithm the training set and the test set are re-sample 10 times and each times the performances of all the algorithms are calculated. The training set is composed by all data related to 12 tools while data of 3 tools are used as test set. In this way, we tested the ability of the algorithms to predicting wear of new tools. The linear model is performed using the classical module *lm* in R. The BLR, instead, is performed using the function $stan_glm$ with the gaussian family. For the decision random forest method, the library *randomForest* is used as a regression module with the default number of 500 trees. For the BLR the method *trainControl* is used with a number of re-sampling equal to 20 and with a linear regression method. The same function is used for the neural network to optimize the training activity of the function *nnet*: the method for the resampling is *cv* for the cross-validation and the number of re-sampling is 4, while in the function *nnet* the parameters are $size = 1$, $decay = 0.01$ and $maxit = 2000$.

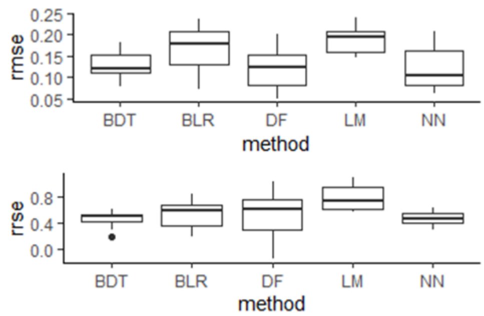

Fig. 4.14 Performances of ML methods according to Root-Mean Square Error and Root Relative Squared Error.

Figure 4.14 reports the results obtained by the 5 algorithms analysed. NN outperformed the other methods by obtaining the lowest error rates with a simple structure with 10 hidden nodes for each of the 3 layers. Figure 7 shows the comparison between the real data and the data estimated by the NN for the tools 14, 15 and 16. Figure 4.15

Among different ML algorithms, Neural Network (NN) was the one with the best performance. For this reason, this model was chosen as the data-driven method to be used in the hybrid model. In Figure 4.16, the weights obtained with a training set composed by 10 tools are represented (ω values). In the later runs, the weights reflect the results obtained with the error analysis, in which there is an intermediate

Fig. 4.15 Comparison between test data and estimated values by NN for wear

linear phase (from run 13 to run 19) in which the Taylor model is more accurate and therefore the weights tend to 1, while in the outer phases the data-driven model based on the NN algorithm prevails.

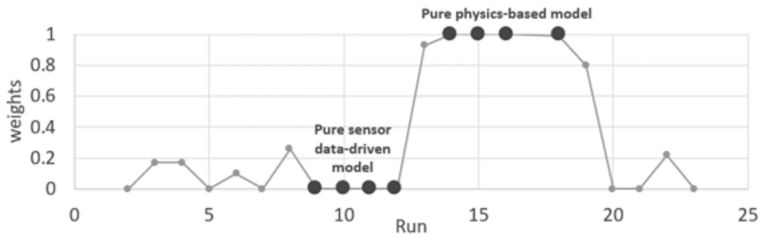

Fig. 4.16 Weights obtained on the training set.

4.4 Results and future implementation

4.4.1 Discussion of results

The results show that the proposed hybrid approach, defined by the linear combination of a physics-based and a data-driven method, has the best performance than either single method. When single models achieve similar performance, the hybrid approach allows to significantly increase the overall accuracy and specially to obtain much more robust results. In addition, this approach can be used as an unique model to estimate tools that are monitored both offline of along each operation.

Figure 4.17 shows the overall errors of the hybrid model and the single models. In terms of the RMSE, the hybrid model has a similar distribution to the NN model, slightly shifted downward and with one less outlier. While analyzing the distribution of the RRSE, the distribution of the hybrid model has a median value similar to the single models but with much less dispersion. It can be inferred that the hybrid model only slightly improves the overall accuracy of the predictions, while the main ad-vantage of this approach is to obtain a more robust method whose goodness of predictions is less affected by the training data than the single models.

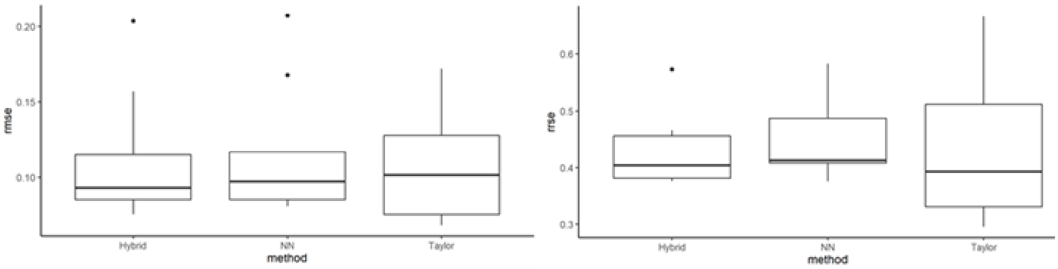

Fig. 4.17 Hybrid model performance compared to individual model performance.

As shown in Figure 4.18, the performance of the hybrid and related single models was analyzed as the size of the training set varied. The error measured with the RRSE metric is always smaller with the hybrid model than with the single models, although the differences are more pronounced with larger training sets. While, analyzing the values of the metric RMSE, it was found that with a training set of small size (composed of 8 tools over total of 14) is more accurate the Taylor model because there are not enough data to train the neural network, which has a much higher error than the physics-based method, and then the hybrid model has intermediate performance between them. While increasing the size of the training set the error of the NN model becomes very similar to that of the Taylor model and consequently the hybrid model obtains better performances than the single models. Moreover, it can be ob-served that with the hybrid model the increase in accuracy with the increase of the tools used in the training set is greater than that obtained by the single models.

The results obtained is a primary formulation of a hybrid model applied on a real manufacturing case study, with the aim of monitoring the status of a CNC machine tool through a combination of an NN model with the extended Taylor law. Considering the two approaches applied on the case study, the physics-based methods result the most accurate and robust, in fact, the best overall prediction performance was obtained with the Taylor model. It has the advantage that it can be implemented even in the absence of sensors on board the machine and with few offline measurements to train the model. The main limitation of this method is that it can only be applied to wear phenomena for which a mathematical law describing

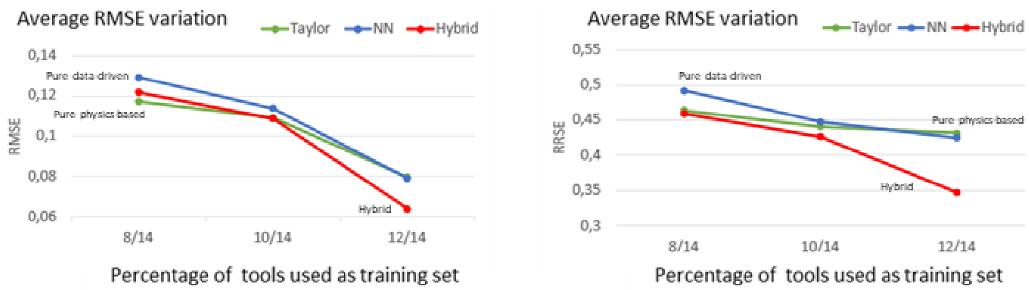

Fig. 4.18 Performance comparison of models for different training set size.

the trend is known, i.e., only for common wear metrics. The category of data-driven methods based on NN model shows it potential during the last milling runs, i.e., with enough sensor data. This characteristic reflects the ability of such methods to explain high variable trends using the information provided by sensors. Limitations of data-driven methods are the need of installing many sensors on the machines for real-time monitoring, and then the high initial investment.

4.4.2 Future research on CNC machine TCM

Despite the theoretic applicability of the framework, on important limitation of this methodology is that the results strongly depend on the amount of data used to train the model. This fact also affects the performance results of the model. This limitation can be solved by attempting to collect data from different machines, by varying a different number of parameters, and add more types of sensors and data sources (e.g., noise, temperature, etc.).

Another limitation consists of the measure of only one flank wear coefficient. Even if is the most commonly parameter used to evaluate the tool wear, other types of wear can occur (e.g., built-up edge, plastic deformation, chipping, notch wear, and thermal cracks). Considering that surely different wear measurements are mutually dependent variables, having a dataset to set up a multivariate problem may demonstrate an improvement in predictive performance as well as a better description of the tool condition given by different types of wear measurements.

The first future improvement regards the dataset. As pointed out by other previous works, the main problem in these kinds of works is the lack of data coming from real environments. Thus, more types of tools need to be considered, with different production parameters and different tool paths depending on the part program. Another improvement of the dataset will consider the use of more sensors, even with different data format, e.g., thermal camera and others. [178] [39] A larger set of sensors would allow an additional analysis of the search for the optimal sensor set by also considering the technology costs during the feature selection phase.

The second future direction is the extension of the framework improving and testing the maintenance strategy and using the results to improve the tool warehouse management and the supply planning. The goal is to optimize the production planning and the maintenance activity, [202] [203] by also considering batch production, [204] and minimizing the energy consumption. [205] [206]

A further direction regards the improvements of the data-driven model. In fact, all the previous features related to the same tool can be used to predict the actual wear level. This means considering a set of features with time depending sizes and this can be managed by extract trends or weighted statistics or developing an algorithm able to receive different number of inputs. Another idea is to use prior defined or fitted distributions for the manufacturing window selection. Also a quantitative evaluation about the size of the training set necessary to obtain acceptable performance could be carried out. Finally, the ML approach that estimate the target variable W_{ML}, used in the hybrid system, can be considered as an hybrid systems powered by all the supervised models shown in this work and not by only the best.

With further research, the general framework can be validated on other case studies, even on different manufacturing processes and with different wear metrics. For example, a proposal for a future work is considering another process completely different from milling, such as welding, and perform experiments to see if the prediction models achieve better performance when trained on the same standardized database, toward a single wear prediction model for different CNC machines. In turn, the hybrid model can be extended to predict the RUL of each tool after each run, in term of remaining runs or estimating the remaining time of usage. In addition, it is possible to implement a hybrid approach between a time series method (starting with a simple auto-regressive one) and the two used in this case study. Other improvements can be done to the hybrid model, such as considering other estimations as input nodes of the NN model.

Finally, a quick analysis work can be devoted to finding a case study that does not lose generality but provides shorter wear times so that we can focus more on validating the framework applied to different machines than on optimizing a single model validated on data that are too homogeneous because with a small in size and few subjects.

Concluding, wear processes are slow processes, and tool breakage events are defined as rare. For this reason, it is expensive to have a dataset that can test the validity of the framework applied to an entire production plant. An experimental plan that is relatively inexpensive but would still provide an acceptable validation of the framework consists of a series of experiments performed on a machine tool (e.g., milling, welding or additive manufacturing machinery) following these requirements:

- consider as a single experiment the execution of a standard processing cycle to be the same for the entire experimental plan,
- use even one type of tool,
- ensure that the number of tool is sufficient for the experimental plan and in any case that it is not less than 30 tools (dataset dimension)
- use at least 2 different materials for the component to be processed,
- consider at least 2 values for each machine setting parameter (tool speed, depth of cut or laser power) included in the physics-based (experimental) model implied in knowledge levels,
- record energy and resource consumption in general both to improve the predictive power of data-driven models and to be able to make economic and environmental impact assessments,
- ensure that the number of tool is sufficient for the experimental plan and in any case that it is not less than 30 tools,
- record the information provided by an expert (prior human knowledge) regarding the wear state of and the optimal replacement time of the tool (probably discretizing these two target variables in order to obtain ordered classes of wear and information about the "Remaining Useful Operation" or the "Remaining Useful Cycle",
- as control sample, measure tool wear levels after the first machine cycle and at the same time as the expert reports the need for tool replacement,

- measuring tool wear levels periodically, trying to use the tool as much as possible, and certainly continuing beyond the expert's replacement report (under the proper safety protocols for the tool)
- product quality evaluation

Chapter 5

Conclusions

5.1 Summary of the work

The main motivation of the research is the desire to investigate hybrid systems and their application in the manufacturing sector. This thesis aims to propose a method or rather a framework for designing an information system that not only includes, but integrates different modeling techniques in order to optimize the stages of mapping the information flows of the system and describing the knowledge needed in order to achieve the awareness that can guide toward optimal choices. Such hybrid systems, i.e., hybrid model systems, turn out to be an interesting research topic as they are widely used even though the scientific literature provides neither a clear definition and comprehensive results regarding their benefits over single models and under different hybridization choices. Another important open research question specifically concerns system knowledge and, in particular, how to integrate so-called prior knowledge and how to structure a framework that provides for the addition of new knowledge sources over time.

The main scientific contribution that this work seeks to bring, therefore, is a proposal for a design method that promotes to the reader, or accentuates, hybrid thinking, that is, designing an information system by considering different models, separately and simultaneously, in order to obtain more reliable descriptions and predictions of the state of the system, ensuring greater resilience of the system as it is able to exploit the strengths of different prediction models. This certainly pushes research efforts toward the concepts of Enterprise Information Systems (EISs) and Knowledge-Based Systems (KBSs), and thus toward the study of how to use different information mapping techniques and different approaches of variables modeling (the main ones being data-driven and those based on laws of physics, analyzing how integrate human-driven ones, i.e., variables models based completely on manufacturing human know-how).

Hybrid modeling, considering different modeling techniques, aim to avoid the unscrupulous use of one single family of models that are often non-sustainable. It is unequivocal that the sustainable manufacturing needs an appropriate digital information system that has (or that is design with) an awareness of the enterprise's objectives and the impact of its use by the enterprise's resources. Today it is required that this awareness is increasingly comprehensive and effective, in other words, that it follows the 5.0 vision by making use (or being able to make use) of all 4.0 technologies. In order to guarantee the sustainability, an hybrid model has to consider the impact due to energy consumption or hardware production and installation, but often the non-sustainability is also due to maintenance costs or the inability of humans to use the system.

In detail, this thesis is a work dealing with (i) awareness Knowledge-Based Systems (KBSs) and (ii) hybrid systems as base concepts for designing a digital platform aiming of supporting a specific physical manufacturing environment. The model proposed by this work is based on these two concepts and, basically, it is a theoretical formalization of a digital platform belonging to a Cyber-Physical System (CPS) that uses hybrid models in order to achieve a 5.0 wisdom, i.e., in order to promote and follow a 5.0 awareness as digital component of information management for a generic 4.0 manufacturing system. The hierarchical structure of the 4 levels of Data, Information, Knowledge and Wisdom (DIKW) is proposed as a method capable of (i) making use of data-lakes, information flows and knowledge processes to make conscious decisions, with wisdom for the precisely, and (ii) developing the concept of hybrid system by characterizing the 4 hybrid subsystems: hybrid data sources, DB and computer networks hybrid models, hybrid models for estimating state variables, and hybrid decision support methods.

The first chapter introduces the background of the author, the motivations and the aim of the research and the research question. This introduction allows, in a few pages, to understand the objective of the thesis and the associated research axes.

The second chapter describes the state of the art carried out starting from lean thinking and waste elimination, trough the digital era of Industry 4.0, until today with the need to have and greater spread of awareness of these tools and their impact. This chapter is a quick introduction to the main manufacturing concepts on which this thesis is based: lean manufacturing from which is important to understand variables and methodologies to eliminate wastes, Enterprise Information Systems (EISs) focusing on ERP, PLM, MES and their integration, Industry 4.0 and its technologies, and finally the Industry 5.0 and the European manufacturing view.

The central chapter presents the methodological proposal in detail and it is a description of the proposed hybrid wisdom-based framework. It includes an analysis of manufacturing processes and the opportunities they offer for the proposed framework. After giving basic notions about agent-based system, DIKW-structures and hybrid modeling, the element of the framework, the agent, is presented with DIKW levels, the hybrid structure of which is subsequently discussed level by level. Finally, a proposal of framework evaluation metrics and the expected impacts in Industry 4.0 and Industry 5.0 is provided.

One case study is presented, and it is focused on the application of the framework to design a Tool Condition Monitoring (TCM) system for a Total Productive Maintenance (TPM). Results obtain with a real case application are provided on milling process: monitoring and optimize the changeover of the milling cutters using open data for results replication. The aims of the case study is to present a design methodology for predictive maintenance functionalities in a TPM CPS.

The theoretical contributions of this work are towards the concepts of hybrid systems, Industry 5.0 and DIKW structures for AI. Even if the single case study presented unfortunately fails to validate all the characteristics required for the general framework, it certainly contributes to various open points of research on hybrid systems. Considering the 6 categories to better integrate social and environmental European priorities into technological innovation and to shift the focus from individual technologies to a systemic approach, this work contributes for the most of them: only bio-inspired technologies and smart materials are difficult to apply for the proposed framework. The proposed work, therefore, can be considered with regard to:

1. individualising Human-Machine Interaction, as agents in CPS modeled as a HW-system trough the proposed framework;

2. digital twins and simulation, as knowledge tools to be integrated with others as components of the hybrid system;

3. data transmission, storage, and analysis technologies described in detail trough the DIKW structure of the proposed framework;

4. Artificial Intelligence structured as MAS where the intelligence is a cognitive back-warding process from data, trough information and knowledge, until the wisdom making use at every level of hybrid modeling systems;

5. technologies for energy efficiency, renewable, storage and autonomy that are considered by the 5.0 oriented wisdom that the framework guides to build.

5.2 Conclusive remarks

The theoretical contributions of this work are towards the concepts of hybrid systems, Industry 5.0 and DIKW structures for AI. For the DIKW method, the section regarding the wisdom characterization of the agent in the third chapter gives a description of the processes in each level and the processes that transform elements lower in the hierarchy into those above them. The next section of the same chapter, on the other hand, provides a definition of hybrid systems for each DIKW level and a proposal to integrate human prior knowledge, simple or dominated by equations. In the same section, the incorporation of the various sources of knowledge is treat.

Even if the single case study presented unfortunately fails to validate all the characteristics required for the general framework, it certainly contributes to various open points of research on hybrid systems:

- the results of the case study are an experimental demonstration about benefits given by hybrid modeling respect using single model,
- a dynamic model selection method is provided (the method of averaging weights varying over time),
- performance of the hybrid model are compared to the ones of the single model,
- other results on the milling open case study are reported in order to promote this dataset for future works,

- data-driven and mechanistic models are used in parallel so that one could help the other.

The results show that the proposed hybrid approach, defined by the linear combination of a physics-based and a data-driven method, has the best performance than either single method. When single models achieve similar performance, the hybrid approach allows to significantly increase the overall accuracy and specially to obtain much more robust results. In addition, this approach can be used as an unique model to estimate tools that are monitored both offline of along each operation.

The literature regarding Industry 5.0 reports other central concepts for this revolution and, even in this case, except for Smart Additive Manufacturing (SAM) which is treated in future improvements, this work shows results along these directions, or more precisely:

- experimental results are taken from the case study on Predictive Maintenance, which is a key manufacturing sector for sustainability and resilience of a system,
- the hyper customization is ensured by the use of product- and service-referenced agents, such that each individual factory resource has a unique correspondence to a DB (and to the entire agent referring to it) or set of them that characterizes it uniquely within the entire CPS,
- cognitive capabilities are assigned to the CPS trough knowledge and wisdom levels toward a Cyber Physical Cognitive Systems (CPCSs), where learning and knowledge are the primary components of decision making that is also at the base of human-robot collaborative manufacturing.

Applying the theoretical framework on the presented case study, several dependencies between the agents were made explicit in defining the wisdom levels: for example, how the measurement agent inherits part of wisdom from the sensor agent, which in turn inherits from the machine driven by the MES through the activity agent. Other relationships not made explicit in the case study are hypothesized and shown in Figure 5.1. From these dependencies, a flow of wisdom generated by source agents directly referring to human beings has been hypothesized. The Figure 5.2 shows this flow and highlights the source agents, i.e., agents in the HW-system directly related to the factory humans and whose wisdom is derived from designing the needs of

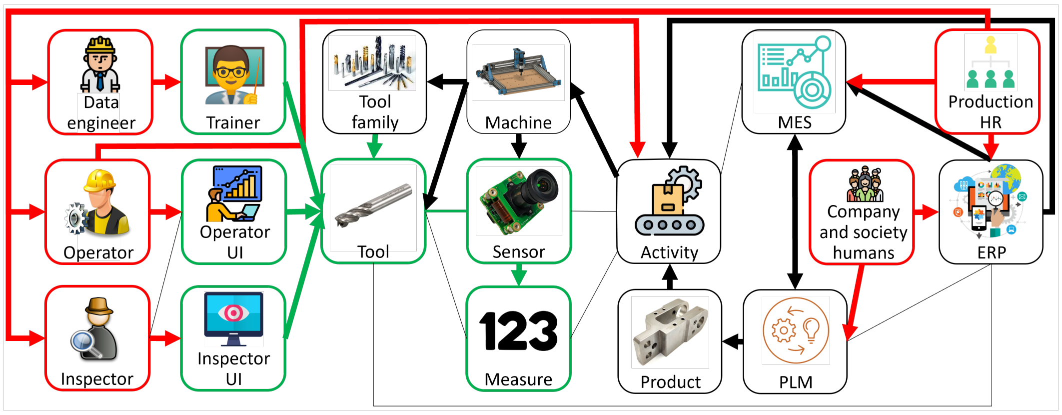

Fig. 5.1 Wisdom relationships estimation of the case study for TCM of CNC machine tool. The simple link between two agents expresses a directional logical dependence, while the arrow indicates the direction by which wisdom is spread (although in reality this is a pull, not a push, process). Red refers to agents strongly associated with humans and the flows by which they spread their wisdom to adjacent agents; green refers to principal, or better characterized in depth, in the case study referring to a machine tool.

society, characterizing of the needs of the humans represented by an operative agent in the system, and keeping going the activity of monitoring the agent's performance and updating the characteristics by interacting whenever possible (and necessary) with the human being. During the design phase of a HW-system, the wisdom levels of each HW-agent are defined through a forward process starting from master agents (sources of wisdom), where the wisdom of dependent agents is defined probably by relaxing some constraints of the inherited wisdom and finding the balance of requirements from the different wisdom sources. In contrast, during the use of CPS in the factory activity, through a backward process, the state of the system is evaluated by focusing on the performance of the wisdom source agents which depend on the performance of agents operating in the system, considering the weight a source agent has in defining the wisdom of the inherited agents.

This wisdom flow mapping methodology allows us to estimate the level of transversality of wisdom that is intended to be provided to the CPS and what are the agents through which this wisdom is digitized. The flow hypothesized in Figure 5.2, while further demonstrating the centrality of the human being in the proposed framework, is a topic for future research as different types of wisdom streams need to be added to study the universal wisdom derived from the design phase and years of CPS deployment and updating the wisdom levels of agents: for example, it is interesting to map the flows related to wisdom generated from (i) technical

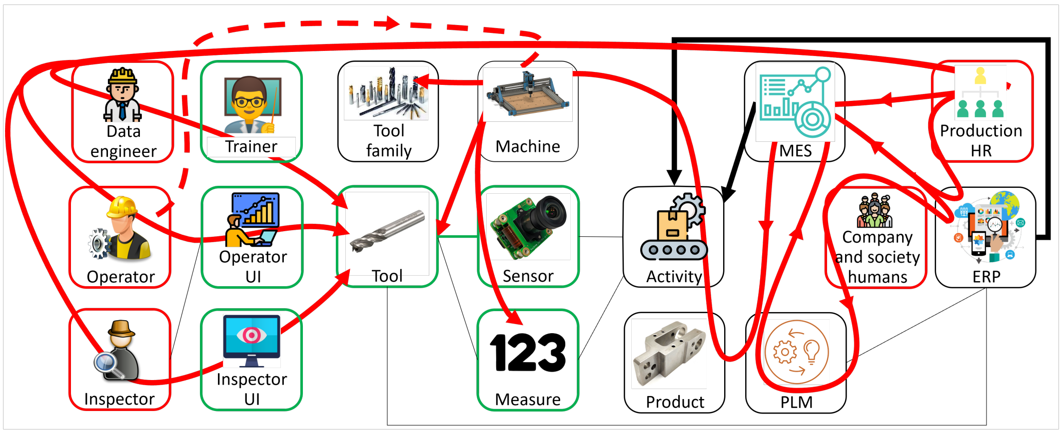

Fig. 5.2 Wisdom flow estimation of the case study for TCM of CNC tool machine. The simple link between two agents expresses a directional logical dependence, while the arrow indicates the direction by which wisdom is spread (although in reality this is a pull, not a push, process). Red refers to agents strongly associated with humans (and the flows by which they spread their wisdom to adjacent agents; green refers to principal, or better characterized in depth, in the case study referring to a machine tool.

documentation and machine use standards, (ii) ERP system (economic, financial and fiscal wisdom), (iii) a hypothetical cost and environmental impact estimation system, or (iv) guidelines from networks between suppliers or partners, commercial groups or governments. Of these wisdom streams, it is interesting to monitor developments during the design phase and the test of a HW-CPS in a real plant (achieve a TRL of 5).

Finally, a final research contribution of this paper is the section in the second chapter on integration between EISs. While integration between MESs and ERP systems is an inherent feature in the hierarchical relationship of these two systems, [114] [115] and integration between ERP and PLM activities remains an area of scientific interest. In particular, many realities that base their core business in One-of-a-Kind Production (OKP) [207] urgently need a complete integration between PLM and MES systems, that, in several cases, remains a complex feature caused by the high heterogeneity of MESs in manufacturing realities. Providing guidelines on standardizing such systems or identifying key components on which to map the lifecy-cle of a generic product is critical. [10] [11] [12] [13] [14]

5.3 Discussion of research questions

The first research question reads as follows:

RQ1. Is the DIKW-schema able to support the design of a Decision Support System (DSS) integrated in a 5.0 smart manufacturing context?

Actually, the size of the scientific literature regarding the current Industry 5.0 does not justify a full interest on the subject. In any case, the concepts of sustainability, resilience and human-centrality are clear pillars from which to build a variables system on which decision-making processes are based to achieve a system able of managing changes in the factory and in the system in which such factory operates, thanks to the collaboration with the human resources of the factory, and in order to make the manufacturing CPS sustainable. The literature regarding the Industry 5.0 underlines the need to find a way to incorporate into the design of AI services some sort of overall vision of the enterprise that can guide AI decisions with factors that include the welfare of broader systems, even broader than the enterprise system itself. In this sense, the DIKW pyramid is a proposal along these line, where the wisdom subsystem represents (i) a concept that drives designers to include such universal factors as the target of the service to be developed, and (ii) a general component common to all the services in the factory in order to create a common vision followed by all digital agents, and non-digital agents, who are able to make effective decisions within the factory.

The second research question reads as follows:

RQ2. What can be a formal and comprehensive definition of hybrid systems employed as manufacturing decision support components?

The second research question refers primarily to the definition of hybrid system and both the analysis of the literature and the proposed framework contribute for the answer to such question: in the literature the use of this term is employed in different ways and occasionally even in conflicting ways. The DIKW method is considered to be a 4-level division of a generic hybrid decision support system in which a definition and several considerations of the hybridization processes occurring in each level have been provided. In particular, the definition of hybrid modeling structures are given considering ensemble models of different families for each DIKW level:

different types of data, heterogeneous information sources (like relational, semi-structured or completely no-structured manageable with AI techniques like NLP) and predictive models using the knowledge generated by ML, PB models, simulations and human-centric evaluations.

Moreover, defining a type of CPS refers to how to measure the contribution of the CPS, which are the models to use in the system and which is its structure, i.e., how these models are integrated. The section on KPIs to be used to assess the applicability of the proposed framework, while part of future work, provides a clear description of the line of research to be followed. In the definition of the framework, the function of creating new model families to hybridize and simply integrating them in the framework, generating new types of information, new knowledge process and new wisdom structures, is discussed characterizing each subsystem.

Finally, the case study provides a relevant scientific contribution as regards the hybridization of ML and PB models, and ideas for the insertion of knowledge directly provided by humans (operator or quality managers) and dynamic simulation models. Comprehensive results regarding their benefits over single models are provided.

The third research question reads as follows:

RQ3. Is hybrid modeling a paradigm that standardize the use of human prior knowledge in manufacturing decision support systems? If so, how?

The third question investigates how to integrate human prior knowledge in the framework and how to structure such framework in order that it is able to generate new knowledge and to receive knowledge from new sources over time. Unfortunately, a single case study, and specifically the case study presented, does not guarantee to fully answer this research question. In any case, in the discussion of the case study different considerations refer to how employing such human a-priori knowledge within wear prediction models. A quantitative assessment of how much such knowledge improves predictive maintenance strategies is not provided due to lack of data, however the theoretical framework presented in the third chapter devotes different parts to the discussion of how to place such a-priori knowledge within the hybridization of predictive models and how to estimate the significance of such knowledge relative to that obtained from mathematical, physical, and statistical models. Finally, the framework also includes a discussion of how such a-priori knowledge, or otherwise not data-driven knowledge, can support the creation of new information structures with the aim of providing new types of knowledge.

5.4 Future works

Future works are presented divided into groups: first the discussion at a general level, then the future works as regards the available data, the third list, the smallest, concerns research in the IT field, while the last two concern models forecasts and the characterization of the wisdom of a CPS.

General improvements and futures activities:

- analyzing different case studies in order to validate the model for all the phases of the entire life cycle of a generic product (and service) or a manufacturing resource;
- defining a standard definition method for the environment agent, i.e., a system containing all the agents and therefore would mainly act as the source of wisdom (it could be considered as a functional supervisor of the main agents in the system, like ERP system, MES, PLM systems or a system given by their integration, or the software environment where the CPS will be developed);
- developing an European research project proposal, starting from the HOME one, where ideas for case studies are (i) electrode condition monitoring and tool changeover management, (ii) MES-PLM integration for AM, (iii) EISs for agriculture management, (iii) Warehouse Management Systems (WMS) and (iv) on-life platform (CPSs) for learning activities, that represent heterogeneous cases, even from outside the manufacture, so as to test the level of generality of the research;
- discussing in depth validation of the framework by comparing it to the hierarchical structure provided by ISA-95;
- developing new models for costing and environment impact evaluation for the agents.

Improvements concerning data and information:

- using case studies with a more complex information flow so as to study the behavior of all agents to assess how they operate;
- link wear levels, or degradation measures in general, with product quality parameters in order to create a system able to determine the optimal changeover strategy considering the desired quality standards;
- working on case studies that are perhaps less scientifically relevant but which present data collection campaigns that are easier to organize also because requiring less IT infrastructure costs (an example is the system for managing the lesson and monitoring student learning).

Knowledge level improvements:

- extending the hybridization methodologies and creating multiple level of hybridization, especially in ML cases, where, instead of selecting the best model, a hybrid structure is always proposed so as to provide better model to be integrated in the first level of hybridization with other models of complete different nature;
- developing a system that can optimize the hybridization parameters in parallel with the choice of ensemble model parameters so that these activities are not left independent;
- exploring different way to hybridize, i.e., other open, or white, box methods, like parallel systems or deterministic in series ones, or grey ones where, for example, the outputs of PB models or the information provided by an operator are considered as additional characteristics to estimate a target variable;
- when it is necessary to estimate the hybridization parameters, using reinforcement learning methods, taking care not to constrain the model to be supervised so as not to lose much of the contribution brought by a-priori knowledge and by the PB models validated in the laboratory.

Wisdom level improvements:

- quantitatively estimating the wisdom flow by starting with defining the wisdom main agents (with different wisdom sources) and defining from them the wisdom of others in cascade following the relationships of the system, i.e., creating a method of estimating the wisdom spread within the factory so as to also have further mapping of the value of company assets based on these concepts;

- collecting values of the indicators used in the wisdom level so as to study their trends, dynamic relationships, and possible drifts of some agents with respect to the system;
- deepen the methods of inheritance of wisdom in order to build a system of replicating agents each with a set of more or less smart functions, but in any case aimed at the business objectives defined by humans so as to place them at the center of all processes of corporate decision making.

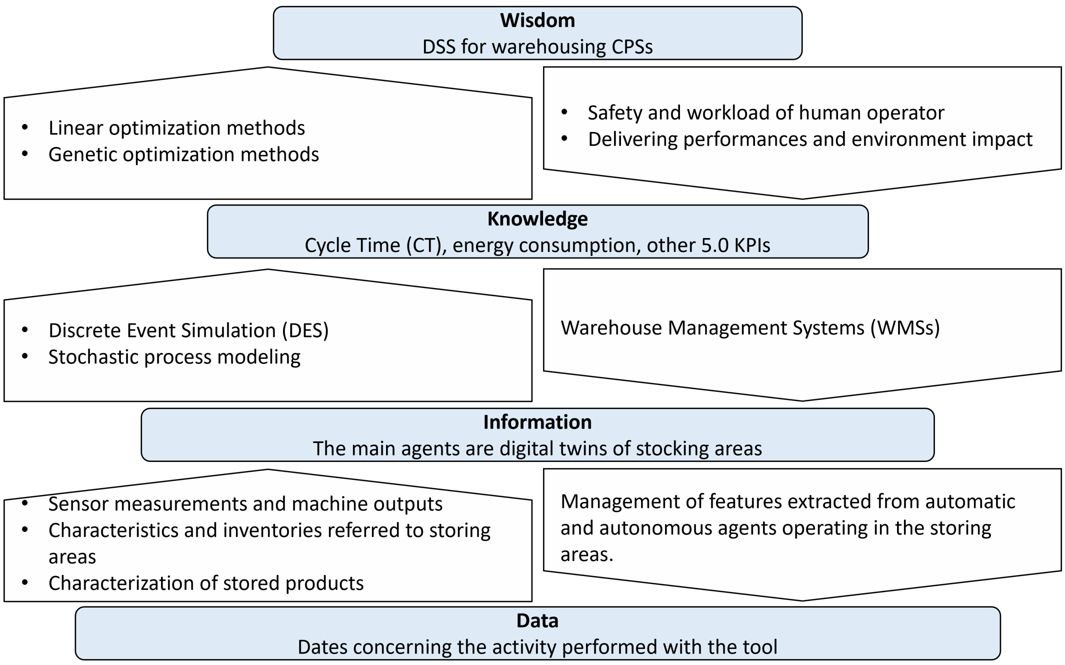

Fig. 5.3 Schema representing the main interesting models to be involved in a future improvements of the project Safe&Green Intralogistic System with 4.0 integrations (SaGrIS4.0) belonging to the MESAP polo.

Finally, starting from the funded research projects outlined in the introductory chapter, interesting future works consist in applying the proposed framework to these three use cases (storage systems, general EISs and AM management) in order to validate it more comprehensively. In particular, as shown by the figure 5.3, from the project SaGrIS4.0 it is possible to plan a second project in order to develop a model for autonomous management of a generic warehouse (including fully automated parts, human-powered parts, and hybrid parts) that looks to the vision 5.0 and especially allows for experimentation with the synergy achieved by hybridizing simulation models, models based on stochastic processes, and, where possible, real-time measurements made directly on the storage facility. The project HOME, instead, is focused on making available the needed information where it is needed, when it is needed, and for who needs it. This makes it extremely complex to ensure the economic, social, environmental, and energy sustainability of manufacturing in sectors that are centuries old (automobile) or more (textile). As shown by the figure 5.4, in order to make manufacturing lean, smart, aware, and sustainable, a future project of the same complexity and importance as the previous one is the perfect environment to validate all the features of the proposed framework. Concluding, the CAPT'N'SEE project is dedicated to professionals willing to enhance their expertise in the use of Additive Manufacturing (AM) technologies. The project provides a scientific contribution toward the dissemination process of manufacturing knowledge (and wisdom) about product design and MES for AM. Therefore, the CAPT'N'SEE project is an excellent starting point for designing the study of how to disseminate that knowledge through an EIS integrated with other generic systems not strictly related to AM: study an AM-specific MES concept that is (i) specific for the management of 3D printers and other AM resources and (ii) able to digitize, generate and share new manufacturing knowledge for AM in the European and global 5.0 industrial context.

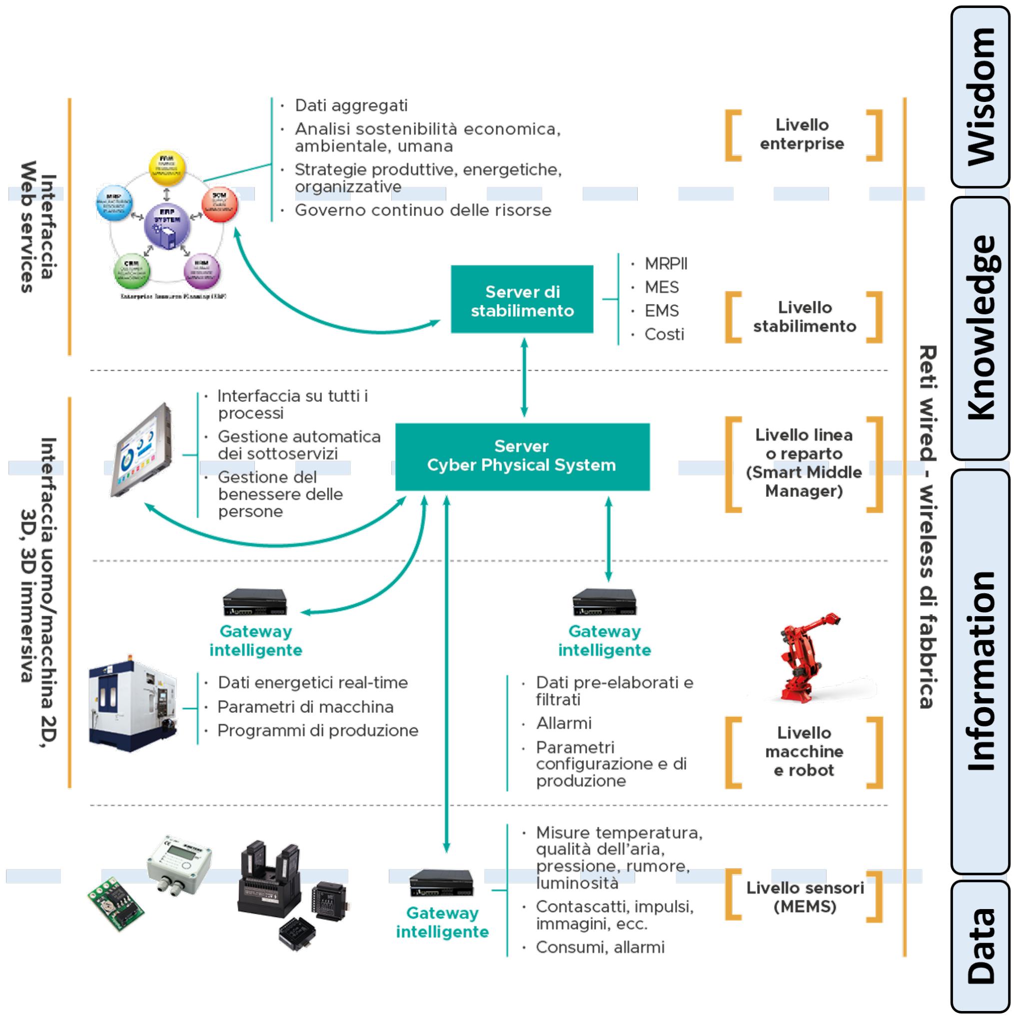

Fig. 5.4 DIKW point of view of the hierarchy of the HOME project, that is funded by the Piedmont Region in the framework of "Programma Operativo Regionale POR-FESR 2014/2020" over the call for tenders of the "Fabbrica Intelligente" technological platform.

References

[1] R. Rai and C.K. Sahu. Driven by data or derived through physics? a review of hybrid physics guided machine learning techniques with cyber-physical system (cps) focus. *IEEE Access*, 2020.

[2] E. Traini, G. Bruno, G. D'Antonio, and F. Lombardi. Machine learning framework for predictive maintenance in milling. *IFAC-PapersOnLine*, pages 177–182, 2019.

[3] E. Traini, G. Bruno, and F. Lombardi. Tool condition monitoring framework for predictive maintenance: a case study on milling process. *International Journal of Production Research*, 2020.

[4] E. Traini, G. Bruno, G. D'Antonio, and F. Lombardi. Design of a physics-based and data-driven hybrid model for predictive maintenance. *APMS 2021: Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems*, pages 536–543, 2021.

[5] L. Panza, G. Bruno, M. De Maddis, F. Lombardi, P. Russo Spena, and E. Traini. Data-driven framework for electrode wear prediction in resistance spot welding. *IFIP Advances in Information and Communication Technology*, 639:239–252, 2022.

[6] G. D'Antonio, G. Bruno, E. Traini, and F. Lombardi. An analytical model to estimate avsrs energy consumption. *IFAC-PapersOnLine*, 52, 2019.

[7] Giulia Bruno, Emiliano Traini, Paolo Chiabert, and Franco Lombardi. Configuration of a production-integrated avsrs through discrete event simulation. *Associazione Italiana Tecnologie Manifatturiere*, 2019.

[8] K. Aliev, E. Traini, M. Asranov, A. Awouda, and P. Chiabert. Prediction and estimation model of energy demand of the amr with cobot for the designed path in automated logistics systems. *Procedia CIRP*, 99:116–121, 2021.

[9] Alberto Faveto, Emiliano Traini, Giulia Bruno, and Franco Lombardi. Development of a key performance indicator framework for automated warehouse systems. *17th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2021*, 54:116–121, 2021.

[10] E. Traini, G. Bruno, A. Awouda, P. Chiabert, and F. Lombardi. Integration between plm and mes for one-of-a-kind production. *IFIP Advances in Information and Communication Technology*, 565:356–365, 2019.

[11] S.A. Venkat, G. Bruno, P. Chiabert, and E. Traini. Integration of plm, mes and erp systems to optimize the engineering, production and business. *IFIP International Conference on Product Lifecycle Management, Springer*, 2019.

[12] G. Bruno, E. Traini, and F. Lombardi. A knowledge-based system for collecting and integrating production information. *IFIP Advances in Information and Communication Technology*, 568:163–170, 2019.

[13] Giulia Bruno, Alberto Faveto, and Emiliano Traini. An open source framework for the storage and reuse of industrial knowledge through the integration of plm and mes. *MANAGEMENT AND PRODUCTION ENGINEERING REVIEW*, 11:62–73, 2020.

[14] Giulia Bruno, Emiliano Traini, Alberto Faveto, and Franco Lombardi. Building a factory knowledge base. *Driving Innovation and Productivity Through Sustainable Automation*, pages 50–75, 2021.

[15] E. Traini, A. Faveto, G. Bruno, and F. Lombardi. Design and simulation of a battery swapping system for electric vehicles. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, 2020-:837–843, 2020.

[16] R. Rosen, Wichert G.V., G. Lo, and K.D. Bettenhausen. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, 48(3):567–572, 2015.

[17] J. Sansana, M.N. Joswiak, I. Castillo, Z. Wang, and R. Rendall. Recent trends on hybrid modeling for industry 4.0. *Computers and Chemical Engineering*, 151:107–365, 2021.

[18] A. Kusiak. Smart manufacturing. *International Journal of Production Research*, 56:508–517, 2018.

[19] F. Tao, Q. Qi, A. Liu, and A. Kusiak. Data-driven smart manufacturing. *Journal of Manufacturing Systems*, 48:157–169, 2018.

[20] X. Yao, J. Zhou, J. Zhang, and C.R. Boer. From intelligent manufacturing to smart manufacturing for industry 4.0 driven by next generation artificial intelligence and further on. *5-th International Conference on Enterprise Systems (ICEIS)*, 48:157–169, 2018.

[21] J.P. Womack, D.T. Jones, and D. Roos. *Toyota Production System: Beyond Large-Scale Production*. Rawson Association, 1988.

[22] V. Gogula, H. Wan, and G. Kuriger. Impact of lean tools on energy consumption. *Sistemas y Telemática*, 9:33, 12 2011.

[23] B. Noori. The critical success factors for successful lean implementation in hospitals. *International Journal of Productivity and Quality Management*, 15:108–126, 01 2015.

[24] R. Gapp, R. Fisher, and K. Kobayashi. Implementing 5s within a japanese context: An integrated management system. *Management Decision*, 46(4):565–579, 2008.

[25] T. Ohno. *The Machine that Changed the World.* CRC Press, 1990.

[26] I.P.S. Ahuja and J.S. Khamba. Tpm for every operator. *Productivity press*, 1988.

[27] I.P.S. Ahuja and J.S. Khamba. Total productive maintenance. *Allied Publishers Limited*, 1988.

[28] I.P.S. Ahuja and J.S. Khamba. Total productive maintenance: literature review and directions. *International Journal of Quality & Reliability Management*, 25(7):709–756, 2008.

[29] T. Dossenbach. Implementing total productive maintenance. *Wood and Wood Products*, 111(2):29–32, 2006.

[30] P. Sangameshwran and Jagannathan. Hll's manufacturing renaissance. *Indian Management,*, pages 30–5, 2002.

[31] F. Ireland and B.G. Dale. A study of total productive maintenance implementation. *Journal of Quality in Maintenance Engineering*, 7(3):277–92, ireland2001.

[32] A. Shamsuddin, M.H. Hassan, and Z. Taha. Tpm can go beyond maintenance: excerpt from a case implementation. *Journal of Quality in Maintenance Engineering*, 11(1):19–42, 2005.

[33] M. Rodrigues and K. Hatakeyama. Analysis of the fall of tpm in companies. *Journal of Materials Processing Technology*, 179(1-3):276–9, 2006.

[34] S. Shingo. *A Study of the Toyota Production System).* Productivity Press, 1989.

[35] S. Shingo. *A revolution in manufacturing: The SMED system.* Productivity Press, 1985.

[36] D.L. Olson and Subodh K. Enterprise information systems: Contemporary trends and issues. *World Scientific*, 2:13–16, 2010.

[37] Gartner Group. A vision of next generation mrp ii. *Tecnical report*, 1990.

[38] M. G. Shields. *E-Business and ERP: Rapid Implementation and Project Planning*. Springer, 2005.

[39] S.I. Chang, G. Gable, E. Smythe, and G. Timbrell. A delphi examination of public sector erp implementation issues. *Proceedings of the Twenty First International Conference on Information Systems*, 7:494–500, 2000.

[40] MESA International. Mes functionalities & mrp to mes data flow possibilities. *Tecnical report*, 1997.

[41] G. D'Antonio, J.S. Bedolla, and P. Chiabert. A novel methodology to integrate manufacturing execution systems with the lean manufacturing approach. *Elsevier - 27th International Conference on Flexible Automation and Intelligent Manufacturing*, 11:2243–2251, 2017.

[42] A. Saaksvuori and A. Immonen. *Product Lifecycle Management*. Springer, 2002.

[43] Stark J. *Product Lifecycle Management*. Springer, 2015.

[44] M. Grives. *Product Lifecycle Management - Driving the next generation of lean thinking*. McGraw-Hill Education, 2005.

[45] R. Sudarsan, S.J. Fenves, R.D. Sriram, and F. Wang. A product information modeling framework for product lifecycle management. *Computer-Aided Design, Elsevier*, 37(13):1399–1411, 2005.

[46] B.R. Leite, M.M. Oliveira, M. Rudek, A. Szejka, and Canciglieri O. Startup definition proposal using product lifecycle management. *Product Lifecycle Management Enabling Smart X, Springer*, pages 426–435, 2020.

[47] M.M. De Oliveira, L.G. Andreatta, J. Stjepandić, and O. Canciglieri. Product lifecycle management and sustainable development in the context of industry 4.0: A systematic literature review. *Transdisciplinary Engineering for Resilience: Responding to System Disruptions*, 16:213–222, 2021.

[48] Y. Zhang, S. Ren, Y. Liu, and S. Si. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of Cleaner Production*, 142:626–41, 2017.

[49] I. Kambarov, G. D'Antonio, K. Aliev, P. Chiabert, and J. Inoyatkhodjaev. Uzbekistan towards industry 4.0. defining the gaps between current manufacturing systems and industry 4.0. *IFIP International Conference on Product Lifecycle Management, Springer*, 48:250–260, 2018.

[50] H. Gaub. Customization of mass-produced parts by combining injection molding and additive manufacturing with industry 4.0 technologies. *Reinforced Plastics*, 60(6):401–404, 2015.

[51] Cynthia Huffman and Barbara E. Kahn. Variety for sale: Mass customization or mass confusion? *Journal of Retailing*, 74(4):491 – 513, 1998.

[52] J.H. Gilmore and B.J. Pine 2nd. The four faces of mass customization. *Harvard business review*, 75(1):91 – 101, 1997.

[53] Giovani Da Silveira, Denis Borenstein, and Flávio S Fogliatto. Mass customization: Literature review and research directions. *International Journal of Production Economics*, 72(1):1 – 13, 2001.

[54] Thomas R. Gruber. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human - Computer Studies*, 43(5-6):907 – 928, 1995.

[55] C. Wittengerg. Cause the trend industry 4.0 in the automated industry to new requirements on user interfaces? *Lecture Notes in Computer Science*, 9171:238–245, 2015.

[56] S. Wang, G. Wang, D. Zhang, D. Li, and C. Zhang. Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. *Computational Network*, 000:1–11, 2016.

[57] B. Bagheri, S. Yang, H. Kao, and J. Lee. Cyber-physical systems architecture for self-aware machines in industry 4.0 environment. *IFAC-Pap Online*, 48(3):1622–1627, 2015.

[58] J. Lee, B. Bagheri, and H. Kao. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3:18–23, 2015.

[59] JL. Thames and D. Schaefer. Cybersecurity for industry 4.0 and advanced manufacturing environments with ensemble intelligence. *Cybersecurity for Industry 4.0*, pages 243–265, 2017.

[60] JL. Thames and D. Schaefer. Industry 4.0: an overview of key benefits, technologies, and challenges. *Cybersecurity for Industry 4.0*, pages 1–33, 2017.

[61] M. Rüß mann, M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, and M. Harnisch. Industry 4.0: the future of productivity and growth in manufacturing industries. *BCG Perspectives*, 2015.

[62] V. Paelke. Augmented reality in the smart factory supporting workers in an industry 4.0 environment. *IEEE Emerging Technology and Factory Automation (ETFA)*, 2014.

[63] A. Syberfeldt, M. Holm, O. Danielsson, L. Wang, and R.L. Brewster. Support systems on the industrial shop-floors of the future—operators' perspective on augmented reality. *Procedia CIRP*, 44:108–113, 2016.

[64] W. Kühn. Digital factory—simulation enhancing the product and production engineering process. *Proceedings of the 2006 winter simulation conference*, pages 1899–1906, 2006.

[65] Y. Zhang, S. Ren, Y. Liu, and S. Si. Big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of Cleaner Production*, 142(2):1–16, 2016.

[66] M. Blanchet, T. Rinn, G. von Thaden, and G. De Thieulloy. Industry 4.0: the new industrial revolution. how europe will succeed. *Think Act, Roland Berger Strategy Consultants GmbH*, 2014.

[67] M. Obitko and V. Jirkovský. Big data semantics in industry 4.0. *International Conference on Industrial Applications of Holonic and Multi-Agent Systems*, 2015.

[68] L. Pittaway, M. Robertson, K. Munir, D. Denyer, and A. Neely. Networking and innova tion: a systematic review of the evidence. *International Journal of Management*, 5:137–168, 2004.

[69] D. Uckelmann. A definition approach to smart logistics. *Lecture Notes in Computer Science*, 5174, 2008.

[70] N. Jazdi. Cyber physical systems in the context of industry 4.0. *IEEE international conference on automation, quality and testing, robotics*, pages 1–4, 2014.

[71] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui. Digital twin-driven product design, manufacturing and service with big data. *International Journal Adv Manufacturing Technology*, 2017.

[72] F. Tao, M. Zhang, J. Cheng, and Q. Qi. Digital twin workshop: a new paradigm for future workshop. *Computer Integrated Manufacturing Systems*, 23(1):1–9, 2017.

[73] M. S. Chair, M. Conroy, R Doyle, Ed Glaessgen, C. Kemp, J. LeMoigne, and L. Wang. Nasa technology roadmap. draft modeling, simulation, information technology & processing roadmap technology area 11. 2010.

[74] M. S. Chair, M. Conroy, R Doyle, Ed Glaessgen, C. Kemp, J. LeMoigne, and L. Wang. Nasa technology roadmap. modeling, simulation, information technology & processing roadmap technology area 11. 2012.

[75] S. Mi, Y. Feng, H. Zheng, Y. Wang, Y. Gao, and J. Tan. Prediction maintenance integrated decision-making approach supported by digital twin-driven cooperative awareness and interconnection framework. *Journal of Manufacturing Systems*, 58:329–345, 2021.

[76] X. Zhang, L. Zhang, and A. Ullah. Digital twin technology and wisdom museum from the perspective of artificial intelligence. *Lecture Notes on Data Engineering and Communications Technologies*, 136:674 – 682, 2022.

[77] B.J. Oakes, B. Meyers, D. Janssens, and H. Vangheluwe. Structuring and accessing knowledge for historical and streaming digital twins. *CEUR Workshop Proceedings*, 2941, 2021.

[78] J. Müller. Enabling technologies for industry 5.0 : results of a workshop with europe's technology leaders. *European Commission and Directorate-General for Research and Innovation*, 2020.

[79] M. Breque, L. De Nul, and A. Petridis. Industry 5.0 : towards a sustainable, human-centric and resilient european industry. *European Commission and Directorate-General for Research and Innovation*, 2021.

[80] A. Renda, S. Schwaag Serger, D. Tataj, A. Morlet, D. Isaksson, F. Martins, M. Mir Roca, C. Hidalgo, A. Huang, S. Dixson-Declève, P. Balland, F. Bria, C. Charveriat, K. Dunlop, and E. Giovannini. Industry 5.0, a transformative vision for europe : governing systemic transformations towards a sustainable industry. *European Commission and Directorate-General for Research and Innovation*, 2022.

[81] M. Maier, A. Ebrahimzadeh, A. Beniiche, and S. Rostami. The art of 6g (tao 6g): How to wire society 5.0. *Journal of Optical Communications and Networking*, 14(2):101–112, 2022.

[82] V. Özdemir, S. Springer, A. Ylldlrlm, S. Biçer, A. Kendirci, S. Sardaş, H. Klllç, N. Hekim, T. Kunej, K.Y. Arga, K. Dzobo, W. Wang, M. Geanta, A. Brand, and M. Bayram. Thanatechnology and the living dead: New concepts in digital transformation and human-computer interaction. *OMICS A Journal of Integrative Biology*, 25(7):401–407, 2021.

[83] S.A. Bhat, N.-F. Huang, I.B. Sofi, and M. Sultan. Agriculture-food supply chain management based on blockchain and iot: A narrative on enterprise blockchain interoperability. *Agriculture (Switzerland)*, 12(1), 2022.

[84] A. Kaklauskas, N. Lepkova, S. Raslanas, I. Vetloviene, V. Milevicius, and J. Sepliakov. Covid-19 and green housing: A review of relevant literature. *Energies*, 14(8), 2021.

[85] G. Paul, N.D. Abele, and K. Kluth. A review and qualitative meta-analysis of digital human modeling and cyber-physical-systems in ergonomics 4.0. *IISE Transactions on Occupational Ergonomics and Human Factors*, 2021.

[86] V.V. Estrela, J. Hemanth, O. Saotome, G. Nikolakopoulos, and R. Sabatini. Imaging and sensing for unmanned aircraft systems. volume 1: Control and performance. volume 2: Deployment and applications. *Imaging and Sensing for Unmanned Aircraft Systems: Control and Performance*, pages 1–346, 2020.

[87] V. Ozdemir and N. Hekim. Birth of industry 5.0: Making sense of big data with artificial intelligence, “the internet of things” and next-generation technology policy. *OMICS A Journal of Integrative Biology*, 22(1), 2018.

[88] S. Nahavandi. Industry 5.0—a human-centric solution. *Sustainability*, 2019.

[89] K.A. Demir, G. Doven, and B. Sezen. Industry 5.0 and human-robot co-working. *Procedia Computer Science*, 158:688–695, 2019.

[90] P.K.R. Maddikunta, Q. Pham, N. Prabadevi, B.and Deepa, K. Dev, T. R. Gadekallu, R. Ruby, and M. Liyanage. Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, 26, 2021.

[91] M. Bernik. After robots take over our jobs, then what? *Forbes*, 2017.

[92] K.A. Demir, C. Ebru, and E. Meral. Issues in integrating robots into organizations. *The 12th International Scientific Conference on Defense Resources Management in the 21st Century*, 2017.

[93] M. Rada. Industry 5.0 definition. *Linkedin*, 2017.

[94] F. Longo, A. Padovano, and S. Umbrello. Value-oriented and ethical technology engineering in industry 5.0: A human-centric perspective for the design of the factory of the future. *Applied Sciences*, 10:4182, 2020.

[95] B. Friedman and D.G. Hendry. Value sensitive design: Shaping technology with moral imagination. *Mit Press*, 2019.

[96] P.J. Koch, M.K. Amstel, P. Debska, M.A. Thormann, A.J. Tetzlaff, S. Bogh, and D. Chrysostomou. A skill-based robot co-worker for industrial maintenance tasks. *Procedia Manufacturing*, 11:83.90, 2017.

[97] Y.K. Leong, J.H. Tan, K.W. Chew, and P.I. Show. Significance of industry 5.0. *The prospect of industry 5.0 in Biomanufacturing*, pages 1–20, 2020.

[98] N.R. Jennings, K. Sycara, and M. Wooldridge. A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, 1:7–38, 1998.

[99] K.P. Sycara. Multiagent systems. *AI Magazine, 19(2), 79*, 19:72, 1998.

[100] M. Wooldridge and N.R. Jennings. Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10:115–152, 1995.

[101] M. Frické. The knowledge pyramid: a citique of the dikw hierarchy. *Journal of Information Science*, 35(2):131–142, 2009.

[102] J. Rowley. The wisdom hierarchy: representations of the dikw hierarchy. *Journal of Information Science*, 33(2):163–180, 2007.

[103] S. Baškarada and A. Koronis. Data, information, knowledge, wisdom (dikw): A semiotic theoretical and empirical exploration of the hierarchy and its quality dimension. *Australasian Journal of Information Systems*, 18(1), 2013.

[104] K. Fan and Y. Duan. *Purpose Computation-Oriented Modeling and Transformation on DIKW Architecture*. Springer, 2022.

[105] M. Frické. The knowledge pyramid: The dikw hierarchyt. *Knowledge Organization*, 46(1):33–46, 2019.

[106] J. Rowley. From data to wisdom. *Journal of applied systems analysis*, 16:3–9, 1989.

[107] L. Floridi. *The Blackwell Guide to the Philosophy of Computing and Information*. Blackwell Philosophy Guides, 2004.

[108] C.E. Shannon. *Collected papers*. IEEE Computer Science Press, 1993.

[109] A. Musgrave. *Common Sense, Science, and Scepticism: a Historical Introduction to the Theory of Knowledge*. Cambridge University Press, 1993.

[110] H.S. Witsenhausen. Hybrid techniques applied to optimization problems. *AFIPS Proceedings - 1962 Spring Joint Computer Conference, AFIPS 1962*, pages 377–392, 1962.

[111] B.A. Mitchell. Hybrid techniques in statistical analysis i. *Simulation*, 5(2):87–88, 1965.

[112] R.F. Brown and G.C. Goodwin. Hybrid method of state and parameter estimation for use in gradient techniques. *Electronics Letters*, 3(12):537, 1967.

[113] R.T. Chien, B.D. Cunningham, and I.B. Oldham. Hybrid methods for finding roots of a polynomial-with application to bch decoding. *IEEE Transactions on Information Theory*, 15(2):329–335, 1969.

[114] International Society of Automation (ISA). Enterprise-control system integration - part 1: Models and terminology. *AMERICAN NATIONAL STANDARD*, ANSI/ISA-95.00.01, 2010.

[115] International Society of Automation (ISA). Enterprise-control system integration - part 2: Objects and attributes for enterprise-control system integration. *AMERICAN NATIONAL STANDARD*, ANSI/ISA-95.00.01, 2010.

[116] R. Kurzweil. Long live ai. *Forbes*, 2005.

[117] M. Treder. Advance human intelligence. *Responsible Nanotechnology*, 2005.

[118] D. Sjödin, V. Parida, M. Kohtamäki, and J. Wincent. An agile co-creation process for digital servitization: A micro-service innovation approach. *Journal of Business Research*, 112:478–491, 2020.

[119] M. Kohtamäki, V. Parida, P. Oghazi, H. Gebauer, and T. Baines. Digital servitization business models in ecosystems: A theory of the firm. *Journal of Business Research*, 104:380–392, 2019.

[120] L. Panait and S. Like. Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, 11:387–434, 20059.

[121] P. Beynon-Davies. *Information Systems: An Introduction to Informatics in Organisationsn*. Palgrave Macmillan, 2002.

[122] A. Bharadwaj, O.A. El Sawy, P. A. Pavlou, and N. Venkatraman. Digital business strategy: Toward a next generation of insights. *MIS Quarterly: Management Information Systems*, 37(2):471–482, 2013.

[123] A. Jashapara. *Knowledge Management: an Integrated Approach.* FT Prentice Hall, 2005.

[124] L.M. Jessup and J.S. Valacich. *Knowledge Management: an Integrated Approach.* FT Prentice Hall, 2003.

[125] J. Rownley. Where is the wisdom that we have lost in knowledge? *Journal of Documentation*, 62(2):251–270, 2006.

[126] J. Cavazos, J.E.B. Moss, , and M.F.P. O'Boyle. Hybrid optimizations: Which optimization algorithm to use? *Springer-Verlag Berlin Heidelberg*, 3923:124–138, 2006.

[127] National Academy of Sciences. Facilitating multidisciplinary research. *National Academy Press*, 2004.

[128] A. Tolk, A. Harper, and N. Mustafee. Hybrid models as transdisciplinary research enablers. *European Journal of Operational Research*, 291:1075–1090, 2021.

[129] B. Giddings, B. Hopwood, and G. O'brien. Environment, economy and society: Fitting them together into sustainable development. *Sustainable Development*, 10:187–196, 2002.

[130] P.S. Sajja and R. Akerkar. Knowledge-based systems for development. *Advanced Knowledge Based Systems: Model, Applications and Research*, 1:1–11, 2010.

[131] M.S. Reis and G. Gins. Industrial process monitoring in the big dataindustry 4.0 era: from detection, to diagnosis, to prognosis. *Processes*, 5:3, 2017.

[132] G. T. Doran. There's a s.m.a.r.t. way to write management's goals and objectives. *Management Review*, 70:35–36, 1981.

[133] V. Chernenkiy, Y. Gapanyuk, V. Terekhov, G. Revunkov, and Y. Kaganov. The hybrid intelligent information system approach as the basis for cognitive architecture. *Procedia Computer Science, Elsevier*, 145:143–152, 2018.

[134] H. A. Bashir and V. Thomson. Metrics for design projects: A review. design studies. *Design Studies*, 20(3):263–277, 1999.

[135] R.J. Bronikowski. Managing the engineering design function. *Van Nostrand Company, New York*, 1986.

[136] P.V. Norden. Useful tools for project management. *Management of production Penguin Books, Baltimore*, page 71–101, 1970.

[137] A. Griffin. A metrics for measuring product development cycle time. *Journal of Product Innovation Management*, 10:112–125, 1993.

[138] A.T. Bahill and W.L. Chapman. Case studies in system design. *International Symposium and Workshop on Systems Engineering of Computer Based Systems, Piscataway*, pages 43–50, 1995.

[139] E. Fleisch, M. Weinberger, and F. Wortmann. Business models and the internet of things. *Whitepaper of the Bosch Internet of Things and Services Lab*, 2014.

[140] V.K. Jain. *Chapter 1 — Overview of big data. Big data and Hadoop.* Publishing Co Ltd, 2017.

[141] D. Delen and H. Demirkan. Data, information and analytics as services. *Decision Support Systems*, 55(1):359–363, 2013.

[142] R. Soltanpoor and T. Sellis. Prescriptive analytics for big data. *27th Australasian Database Conference: ADC 2016. Springer International Publishing, Sydney, NSW.*, page 245–256, 2016.

[143] D.H. Guston. Responsible innovation: Who could be against that? *Journal of Responsible Innovation*, 2:1–4, 2015.

[144] R.K. Mobley. An introduction to predictive maintenance. *Elsevier*, 2002.

[145] G. Hale. Intech survey: Predictive maintenance top technology challenge in 2007. *InTech*, 2007.

[146] G.P. Sullivan, R. Pugh, A.P. Melendez, and W.D. Hunt. Operations & maintenance best practices - a guide to achieving operational efficiency (release 3.0). *Pacific Northwest National Laboratory*, 2010.

[147] B. Cline, R. S. Niculescu, D. Huffman, and B. Deckel. Predictive maintenance applications for machine learning. *IEEE*, 2017.

[148] L.D. Xu, E.L. Xu, and L. Li. Industry 4.0: state of the art and future trends. *International Journal of Production Research*, 2018.

[149] M.S. Kan, A.C.C. Tan, and J. Mathew. A review on prognostic techniques for non-stationary and non-linear rotating systems. *Mechanical Systems and Signal Processing*, pages 62–63, 2015.

[150] A. Heng, S. Zhang, A.C.C. Tan, and J. Mathew. Rotating machinery prognostics: State of the art, challenges and opportunities. *Elsevier*, pages 724–739, 2009.

[151] H.P. Bloch and F.K. Geitner. Machinery failure analysis and troubleshooting. *Butter-worth-Heinemann*, 2012.

[152] Q. Zhang, P.W. Tse, X. Wan, and G. Xu. Remaining useful life estimation for mechanical systems based on similarity of phase space trajectory. *Elsevier*, pages 2353–2360, 2015.

[153] Y. Lei. An improved exponential model for predicting remaining useful life of roll-ing element bearings. *IEEE Transactions on Industrial Electronics*, pages 7762–7773, 2015.

[154] D. An, N.H. Kim, and J. Choi. Practical options for selecting data-driven or physics-based prognostics. *Reliability Engineering and System Safety*, 133:223–236, 2015.

[155] J. Wang, X. Zhang, J. Zeng, and Y. Zhang. Optimal dynamic imperfect preventive maintenance of wind turbines based on general renewal processes. *International Journal of Production Research*, 58(22):6791–6810, 2019.

[156] L. Xiao, T. Xia, E. Pan, and X. Zhang. Long-term predictive opportunistic replacement optimisation for a small multi-component system using partial condition monitoring data to date. *Zhang International Journal of Production Research*, 58:4015–4032, 2020.

[157] A. Widodo, E. Kim, J.D. Son, and B.S. Yang. Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine. *Expert Systems with Applications*, 3(3):7252–7261, 2009.

[158] W. Caesarendra, A. Widodo, and B.S. Yang. Application of relevance vector machine and logistic regression for machine degradation assessment. *Mechanical System and Signal Processing*, 24(4):1161–1171, 2010.

[159] M.D. Prieto. Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Transactions on Industrial Electronics*, 60(8):3398–3407, 2012.

[160] D.H. Pandya, S.H. Upadhyay, and S.P. Harsha. Fault diagnosis of rolling element bearing by using multinomial logistic regression and wavelet packet transform. *Soft Computing*, 18:255–266, 2014.

[161] P.J.G. Nieto, E. Garcia-Gonzalo, F.S. Lasheras, and F.J. De Cos Juez. Hybrid pso–svm–based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. *Reliability Engineering System Saferty*, 138, 2015.

[162] A. Soualhi, K. Medjaher, and N. Zerhouni. Bearing health monitoring based on hilbert-huang transform, support vector machine and regression. *IEEE Transactions on Industrial Electronics*, 64(1):52–62, 2015.

[163] C. Li, R.V. Sanchez, G. Zurit, M. Cerrada, D. Cabrera, and R.E. Vásquezc. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing*, 168(30):119–127, 2015.

[164] U. Shafi, A.A. Safi, A.R. Shahid, and S. Ziauddin. Vehicle remote health monitoring and prognostic maintenance system. *Journal of Advanced Transportation*, 1:1–10, 2018.

[165] W. Ahmada, S. AliKhana, M.M.M. Islam, and J.M. Kim. A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models. *Reliability Engineering and System Safety*, 184:67–76, 2019.

[166] T. Wu, L. Sari, B.T. Lin, and C.W. Chang. Monitoring of punch failure in micro-piercing process based on vibratory signal and logistic regression. *The International Journal of Advanced Manufacturing Technology volume*, 93:2447–2458, 2017.

[167] H. Li, Y. Wang, P. Zhao, X. Zhang, and P. Zhou. Cutting tool operational reliability prediction based on acoustic emission and logistic regression model. *Journal of Intelligent Manufacturing volume*, 26:923–931, 2015.

[168] D. you, X. Gao, and S. Katayama. Wpd-pca-based laser welding process monitoring and defects diagnosis by using fnn and svm. *IEEE Transactions on Industrial Electronics*, 62(1):628–636, 2015.

[169] F. Jia, Y. Lei, J.L.X. Zhou, and N. Lu. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, 72-73:303–315, 2016.

[170] L. Wang, Z. Zhang, H. Long, J. Xu, and R. Liu. Wind turbine gearbox failure identification with deep neural networks. *IEEE Transactions on Industrial Electronics*, 13(3):9–12, 2017.

[171] A. Krishnakumari, A. Elayaperumal, M. Saravanan, and C. Arvindan. Fault diagnostics of spur gear using decision tree and fuzzy classifier. *The International Journal of Advanced Manufacturing Technology*, 89:9–12, 2017.

[172] D. Wu, J. Jennings, C. andTerpenny, and S. Kumara. Cloud-based machine learning for predictive analytics: Tool wear prediction in milling. *The International Journal of Advanced Manufacturing Technology*, pages 2062–2069, 2017.

[173] Y. Yu. Tool condition prognostics using logistic regression with penalization and manifold regularization. *Applied Soft Computing*, 64:454–467, 2018.

[174] G. Li, H. Chen, Y. Hu, and J. Wang. An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators. *Applied Thermal Engineering*, 129, 2018.

[175] P. Santos, J. Maudes, and A. Bustillo. Identifying maximum imbalance in datasets for fault diagnosis of gearboxes. *Journal of Intelligent Manufacturing*, 29:333–351, 2018.

[176] A. Agogino and K. Goebel. Milling data set. *NASA Ames Prognostics Data Repository. BEST lab, UC Berkeley. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/*, 2007.

[177] L. Panza, M. De Maddis, and P. Russo Spena. Use of electrode displacement signals for electrode degradation assessment in resistance spot welding. *Journal of Manufacturing Processes*, 76:93–105, 2022.

[178] H.M. Hashemian and C.B Wendell. State-of-the-art predictive maintenance techniques. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, 60(10), 2011.

[179] G.J. Ross, D.K. Tasoulis, and N.M. Adams. Nonparametric monitoring of data streams for changes in location and scale. *JTechnometrics*, pages 379–389, 2012.

[180] G.J. Ross and N.M. Adams. Two nonparametric control charts for detecting arbitary distribution changes. *Journal of Quality Technology*, page 102:116, 2012.

[181] M. Haghighat, M. Abdel-Mottaleb, and W. Alhalabi. Recognition, discriminant correlation analysis: Real-time feature level fusion for multimodal biometric. *IEEE Transactions on Information Forensics and Security*, 11(9):1984–1996, 2016.

[182] L. Al Shalabi, Z. Shaaban, and B. Kasasbeh. Data mining: A preprocessing engine. *Journal of Computer Science*, 2:735–739, 2006.

[183] G. Upton and I. Cook. Understanding statistics. *Oxford: Oxford University Press*, 1996.

[184] V. Kumar and S. Minz. Feature selection: A literature review. *Smart Computing Review*, 4(3):211–219, 2014.

[185] E. Alpaydin. Introduction to machine learning. *MIT Press*, 2014.

[186] P. Stoica and R. Moses. Spectral analysis of signals. *Prentice Hall*, 2005.

[187] E. Parzen. Stochastic processes. *Couvier Dover Publications*, 2015.

[188] V. P. Astakhov. Effects of the cutting feed, depth of cut, and workpiece (bore) diameter on the tool wear rate. *The International Journal of Advanced Manufacturing Technology*, pages 631–640, 2007.

[189] M. von Stosch, R. Oliveira, J. Peres, and S. Feyo de Azevedo. Hybrid semi-parametric modeling in process systems engineering: past, present and future. *Computers & Chemical Engineering*, 101:60/86, 2014.

[190] R. P. A. Siddhpura. A review of flank wear prediction methods for tool condition monitoring in a turning process. *International Journal of Advanced Manufacturing Technology*, 65(1-4):371–393, 2012.

[191] G.J. Ross. Sequential and batch change detection using parametric and nonparametric methods. *CRAN*, 2015.

[192] M. Frigo and J. G. Johnson. The design and implementation of fftw3. *IEEE*, pages 216–231, 2005.

[193] W. J. Hopp and M. L. Spearman. Factory physics. *Long Grove: Waveland Press.*, 2011.

[194] J. Coble and J. W. Hines. Identifying optimal prognostic parameters from data: A genetic algorithms approach. *Annual Conference of the Prognostics and Health Management Society*, 2009.

[195] K. Aslantas, I. Ucun, and A Cicek. Tool life and wear mechanism of coated and uncoated al_2o_3/ticn mixed ceramic tools in turning hardened alloy steel. *Wear*, 274-275:442–451, 2012.

[196] L. Dan and J. Mathew. Tool wear and failure monitoring techniques for turning - a review. *International Journal of Machine Tools and Manufacture*, 579-598, 2011.

[197] T. Hasti, R. Tibshirani, and J. H. Friedman. The elements of statistical learning: data mining, inference, and prediction. *New York: Springer*, 2009.

[198] J. V. Abellan-Nebot. A review of machine monitoring systems based on artificial intelligence process models. *International Journal of Advanced Manufacturing*, 47:237–257, 2010.

[199] Z. Liang. A study of tool life sensitivity to cutting speed. *ASEE PEER*, 2002.

[200] A.L.B. Dos Santos, M. Duarte, A.M. Abrao, and A Machado. Rotating machinery prognostics: State of the art, challenges and opportunities.an optimisation procedure to determine the coefficients of the extended taylor's equation in machining. *International Journal of Machine Tools and Manufacture*, pages 17–31, 1999.

[201] J.R.F. Arruda and M. Duarte. Updating rotor-bearing finite element models using experimental frequency response functions. *AIAA Journal*, 1993.

[202] X. Zhou, M. Zhu, and W. Yu. Maintenance scheduling for flexible multistage manufacturing systems with uncertain demands. *International Journal of Production Research.*, 2020.

[203] Y. Zhao, Y. He, D.Z. Zhou, X. Han, Y. Li, and W. Wang. Functional risk-oriented integrated preventive maintenance considering product quality loss for multistate manufacturing systems. *International Journal of Production Research.*, 2020.

[204] J. Huang, L. Wang, and Z. Jiang. A method combining rules with genetic algorithm for minimizing makespan on a batch processing machine with preventive maintenance. *International Journal of Production Research.*, 58(13), 2020.

[205] Z. Hajej and N. Rezg. An optimal integrated lot sizing and maintenance strategy for multi-machines system with energy consumption. *International Journal of Production Research.*, 58(14), 2020.

[206] Z. Hajej, R. Nidhal, C. Anis, and M. Bouzoubaa. An optimal integrated production and maintenance strategy for a multi-wind turbines system. *International Journal of Production Research.*, 2019.

[207] Y. Tu and P. Dean. *One-of-a-Kind Production*. Springer, 2011.

Appendix A

Appendix

A.1 Author's scientific publications

My publications about researches performed during each year of the PhD course in "Management, Production and Design".

- Year 2019

  Conference paper. Machine Learning Framework for Predictive Maintenance in Milling. [2]

  Conference paper. An analytical model to estimate AVSRS energy consumption. [6]

  Conference paper. Integration Between PLM and MES for One-of-a-Kind Production. [10]

  Conference paper. Integration of PLM, MES and ERP Systems to Optimize the Engineering, Production and Business. [11]

  Conference paper. Configuration of a production-integrated AVSRS through discrete event simulation. [7]

  Conference paper. A Knowledge-Based System for Collecting and Integrating Production Information. [12]

- Year 2020

Journal. Tool condition monitoring framework for predictive maintenance: a case study on milling process. [3]

Journal. An open source framework for the storage and reuse of industrial knowledge through the integration of PLM and MES. [13]

Conference paper. Design and Simulation of a Battery Swapping System for Electric Vehicles. [15]

- Year 2021

  Conference paper. Design of a physics-based and data-driven hybrid model for Predictive Maintenance. [4]

  Conference paper. Development of a key performance indicator framework for automated warehouse systems. [9]

  Essay. Building a Factory Knowledge Base. [14]

  Essay. Prediction and estimation model of energy demand of the AMR with cobot for the designed path in automated logistics systems. [8]

- Year 2022

  Conference paper. Data-Driven Framework for Electrode Wear Prediction in Resistance Spot Welding [5]