Journal of Machine Engineering, 2021, Vol. 21, No. , 5–39
ISSN 1895-7595 (Print) ISSN 2391-8071 (Online)

Received: *26 June 2021* / Accepted: *31 October 2021* / Published online: *03 September 2021*

data analytics, process optimization, semantics, Industry 5.0

Dimitris MOURTZIS[1*]

TOWARDS THE 5th INDUSTRIAL REVOLUTION: A LITERATURE REVIEW AND A FRAMEWORK FOR PROCESS OPTIMIZATION BASED ON BIG DATA ANALYTICS AND SEMANTICS

The digitalization of modern manufacturing systems has resulted to increasing data generation, also known as Big Data. Although there are several technologies and techniques under the term Data Analytics for gathering such data, their interpretation to information, and ultimately to knowledge remains in its infancy. Consequently, albeit engineers currently can monitor the factory level, optimization is cut off of the data acquisition, and is based on data related methodologies. The focus should be pivoted on designing and developing suitable frameworks for integrating Big Data to process optimization based on the context of information gathered from the shopfloor. This paper aims to investigate the opportunities and the gaps as well as the challenges arising in the current industrial landscape, towards the efficient utilization of Big Data, for process optimization based on the integration of semantics. To that end, a literature review is performed, and a data-based framework is presented.

1. INTRODUCTION

Manufacturing is one of the most important pillars of any modern economy. In Europe, almost 35 million people were employed in the manufacturing sector in 2020, accounting for 15% of European Gross Domestic Product (GDP) [1]. Labour productivity in the manufacturing sector is expected to increase by nearly 10% by 2025 [2, 3]. Local economies have evolved into global and highly competitive players over the last few decades. Industries began to operate on an intercontinental scale, broadening the scope of their operations. Up until the 1990s, the export of finished goods to foreign markets was the dominant theme in international trade, and it has gotten even more attention in the last decade. Furthermore, specific factors to each location, such as low-cost labour and highly skilled personnel, aided rapid globalization [4]. Furthermore, the advent of the Internet and the increasing computational power led to the creation of virtual entities [5], which transformed competition and collaboration practices [6].

[1] Laboratory for Manufacturing Systems and Automation (LMS), Department of Mechanical Engineering and Aeronautics, University of Patras, Greece
[*] E-mail: mourtzis@lms.mech.upatras.gr
https://doi.org/10.36897/jme/141834

1.1. EVOLUTION OF MANUFACTURING PARADIGMS

Since its birth two centuries ago, the manufacturing industry has gone through a number of paradigm shifts. The most prevalent manufacturing paradigms that have characterized significant periods of time are as follows: a) Craft Production, b) American Production, c) Mass Production, d) Lean Production, e) Mass Customization, and f) Global Manufacturing as illustrated in Fig. 1.

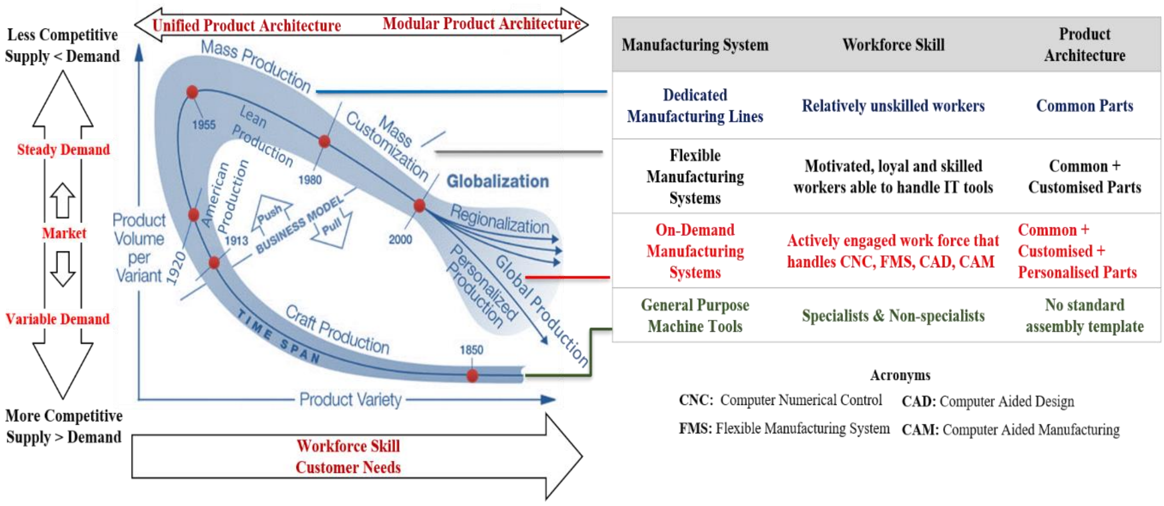


Fig. 1. Manufacturing Paradigm Shifts and Drivers and differences between Production Paradigms [7]

Over the last decade, the shift toward manufacturing digitalization has gotten a lot of attention. Unpredictable demand volatility, higher quality requirements, personalized and customized commodities, and the advent of smart supply chains are the main challenges of modern production systems [8]. Looking back at previous industrial revolutions, the first began at the end of the eighteenth century with the increased use of steam and water power, resulting in a transition from hand production methods to machines (i.e. mechanisation); the second began in the late nineteenth century, utilizing electrical energy and enabling mass production (i.e. intensive use of electrical power); and the third began in the late twentieth century, utilizing electrical energy and enabling mass production (i.e. intensive use of electrical power). Next, from the 1970s, the third made use of electrical and internet technology, as well as automated productions (i.e. digitalisation). We are currently in the fourth industrial revolution, known as Industry 4.0. We have arrived at a stage when there is a clear application-pull (industry demands) and technology-push (technical advancements) working together to propel this new revolution forward. On the one hand, there is a significant demand for change in general social, economic, and political spheres [9]. The digitalization of manufacturing introduces an ecosystem of different technologies in the fields of sensing, connectivity, data modelling and decision-making [10]. Furthermore, communication networks and frameworks such as Wireless Sensor Networks (WSN), Industrial Networks, and Web services are critical enablers of the continuous flow of information between systems [11].

1.2. THE INDUSTRIAL, OPERATOR AND MACHINE TOOL TIMELINE EVOLUTION

The First Industrial Revolution began in the 1780s with the development of mechanical power from water, steam, and fossil fuels. In the 1870s, manufacturers with assembly lines and mass production favoured electrical energy in the second such revolution. In the 1970s, the Third Industrial Revolution introduced the concept of automation to the manufacturing industries using electronics and information technologies (IT). The Internet of Things (IoT) and cloud computing are used in the fourth revolution or Industry 4.0 to provide a real-time interface between the virtual and physical worlds, also known as cyber-physical systems [12]. While it took hundreds of years for the first three industrial revolutions to be happen, the fourth industrial revolution was only coined in 2011. A key characteristic is the rise of digitalisation, which was boosted by technologies like Internet of Things (IoT), cognitive computing, and the combination of Big Data and Artificial Intelligence (AI) [13]. At the same time, while Industry 4.0, with its technological decentralization and interconnectivity, is still in full mode, Industry 5.0, with its full integration of the human touch in business and intelligent systems, will inevitably succeed it. It is expected that the combination of machines and humans working together will merge the increased accuracy and repeatability of full automation with the cognitive skills of the expert managers [14].

The increasing integration of AI creates many challenges and opportunities for the workplaces of the future. Even though Industry 4.0 is still in its infancy, many industry pioneers and technology leaders are anticipating the Fifth Industrial Revolution or Industry 5.0: autonomous manufacturing with human intelligence in and on the loop. To that end, Industry 5.0 combines the two main pillars of Industry 4.0, automation, and efficiency, with a personal human touch. People who work alongside robots, smart machines, and technologies are referred to as co-workers. The evolution of machine tools has been substantially affected by the evolution of industrialization. Industry 1.0 (mechanization, end of the 18th century), Industry 2.0 (mass production, beginning of the 20th century), Industry 3.0 (automation and IT, beginning of the 1970s) and Industry 4.0 (digitalization based on cyber-physical structures, present time) are presented in Fig. 2 [15, 16]. Likewise, the evolution of machine tools is summarized in four stages as follows: Machine Tool 1.0 (mechanically driven but manually operated, late 18th century), Machine Tool 2.0 (electronically driven and numerically controlled, mid-20th century), Machine Tool 3.0 (computer numerically controlled, late 20th century), and Machine Tool 4.0 (computer numerically controlled, late 20th century) (Cyber-Physical Machine Tool 4.0) [17]. The phrase “Machine Tool 4.0” refers to a modern technical development of machine tools that has been fuelled by recent advances in Cyber-Physical Systems (CPS), the Internet of Things (IoT), and Cloud-based applications. In addition, the Operator 4.0 (O4.0) has emerged as a new term in the Industry 4.0 framework, tracing the history of generators alongside the first three industrial revolutions. According to [18], an O4.0 is “a smart and professional operator who performs not only robot cooperative work but also machine-assisted work if and when necessary”. Furthermore, according to [19], the O4.0 is a hybrid agent designed as a symbiotic interaction between the person and the machine, with a focus on treating automation as a further development of the human's physical, sensory, and cognitive capacities.

Manufacturing systems transformed into digital ecosystems as the fourth industrial revolution emerges [13]. The IoT and Big Data play a significant role in this transformation. To that end, industrial companies have entered a new era known as "Big Data era" in which the volume, velocity, and variety of data they manage is rapidly increasing [20]. Big data is a term that refers to a large amount of data that is difficult to process using traditional database techniques. It can be structured or unstructured, and both types offer numerous advantages [20]. In intelligent manufacturing, industrial big data not only enable businesses to accurately perceive changes in the internal and external environments of a system, but also facilitate the integration of advanced scientific analysis and decision-making methodologies, in an attempt to improve the production process, reduce costs, and increase operational efficiency. New business models, such as mass customization and precision marketing are emerging because of massive data, empowering social development, and economic growth [21–24]. As a result, big industrial data is being viewed as a means of driving intelligent manufacturing towards Industry 5.0. Therefore, big data analytics (BDA) has greatly improved with the advancement of AI to effectively mine both structured and unstructured industrial data in intelligent manufacturing [25]. The operation of manufacturing systems will be dramatically changed as BDA continues to develop [26].

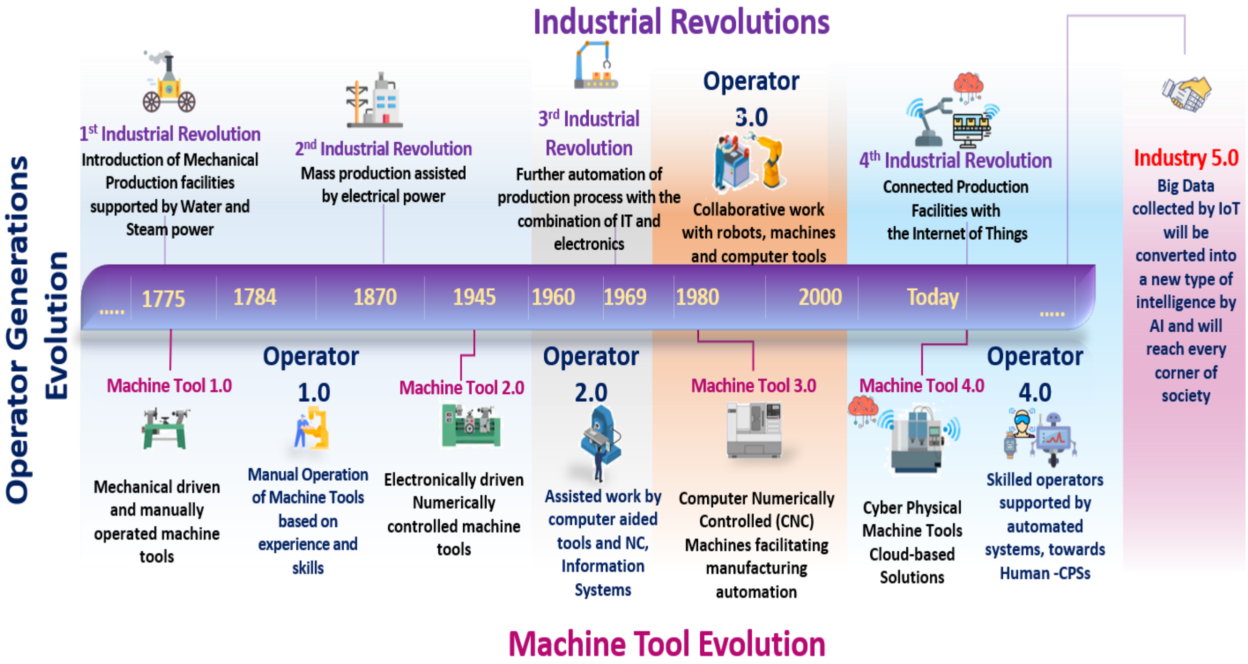


Fig. 2. Evolution of Industry, Machine and Operator [15, 16]

Industry 4.0 reference architectures have emerged to standardize and advance such systems. Industry 4.0 architectures are use case specific, with some common ground in implementation methods. This has increased the demand for standards and guidelines in the field. The main drivers behind the development of Industry 4.0 reference architectures are interoperability, system development simplification, and ease of implementation. There are three major Industry 4.0 reference architectures: Reference Architecture Model Industrie 4.0

(RAMI 4.0), Industrial Internet Reference Architecture (IIRA), and Internet of Things – Architecture (IoT-A). These architectures contain general guidelines for Industry 4.0 architecture design. The RAMI 4.0 model goes beyond pure Industry 4.0 system modelling by incorporating manufacturing and logistics aspects. The IIRA model, on the other hand, may support a higher level of detail in industrial applications. Finally, the IoT-A model places a strong emphasis on the information technology used in an Industry 4.0 application [27]. The RAMI 4.0 model was used to model the current set of tools because it is focused on product development and production scenarios within the context of an Industry 4.0 enterprise. RAMI 4.0 is the result of an effort led by industrial stakeholders to support the global trend of Internet of Things (IoT) research and adoption. It includes a reference architecture for IoT systems that was developed in collaboration with the various organizations involved. It is framed in three dimensions and reflects the fundamental features of the globally accepted Smart Grid Architecture Model [28]. The RAMI 4.0 three-dimensional model is made up of the hierarchy levels axis, the life cycle and value stream axis, and the various layers that form the horizontal axis. Towards addressing the abovementioned challenges, this paper presents a literature review and a framework for process optimization based on Big Data Analytics and Semantics towards the 5th Industrial Revolution (Industry 5.0).

The remainder of the paper is structured as follows. In Section 2, a literature review is presented, in order to highlight the milestones achieved through the years in each industrial revolution, as well as the key improvements identified in manufacturing systems. Further to that, an investigation of the most pertinent research works, in the field of big data and semantics is performed, targeting at the identification of the literature gaps and the key technological aspects that will lead to the fifth industrial revolution. Then in Section 3, the focus is concentrated on the integration of semantics in Big Data Analytics, which by extension will lead to more efficient data interpretation in the manufacturing domain. Afterwards, in Section 4, the opportunities emerging from the integration of Artificial Intelligence in Big Data Analytics are examined. Then, in Section 5, based on the previous paragraphs, a conceptual framework is presented. In Section 6, a short discussion upon the key implementation steps is presented. Finally, Section 7 summarizes the paper and future research direction are discussed.

2. LITERATURE REVIEW

In the following paragraphs, a detailed literature review and discussion is presented. More specifically, the key aspects, and the cornerstone technologies that have already affected and will continue to facilitate the ongoing transformation of the modern manufacturing and production landscape are analysed. More specifically, in Section 2.1 special emphasis is given in Big Data Analytics in parallel with the evolution of AI technologies and techniques. Then in Section 2.2 the light is shed on the key characteristics of modern manufacturing and production systems, which are required for the transition to the next industrial revolution, also known as Industry 5.0.

2.1. BIG DATA STATISTICS AND DEVELOPMENT

Data generated by sensors embedded in machine tools, cloud-based solutions, and business management have already reached a total volume of more than 1000 Exabytes annually in modern industries. According to IDC, data has reached 40 trillion gigabytes in 2020. In 2010, the total amount of big data was 1.2 zettabytes. However, according to the same report, data doubled every two years until 2020 [29]. Additionally, according to an IBM study, the 90% of data was created during 2015 and 2017 [30]. The exponential growth of big data appears unstoppable. The amount of data available on the internet is enormous, and new data is being added every second. Online, 2.5 quintillion bytes of data are generated every day. According to Physics.org, downloading all the data available on the internet would take around 181 million years. The increase in data is unsurprising, given that in 2018, internet users spent 2.8 million years online. Search engines, e-commerce, and social media have all played a role in the global population's high level of internet usage.

Furthermore, advanced analytics can help manufacturers decrease process defects and save time and money. Manufacturers have been able to drastically enhance product quality and yield (the amount of output per unit of input) by introducing lean and Six Sigma programs in their manufacturing processes over the last 20 years or more. Many multinational manufacturers across a wide range of industries now have access to real-time shop-floor data as well as the ability to do complex statistical analyses. They are combining and analysing previously separate data sets in order to uncover crucial insights. Although the big data era is still in its early stages, advanced analytics is based on years of mathematical research and scientific application. It can be a crucial tool for achieving yield gains, especially in industrial environments with high process complexity, variability, and capacity constraints.

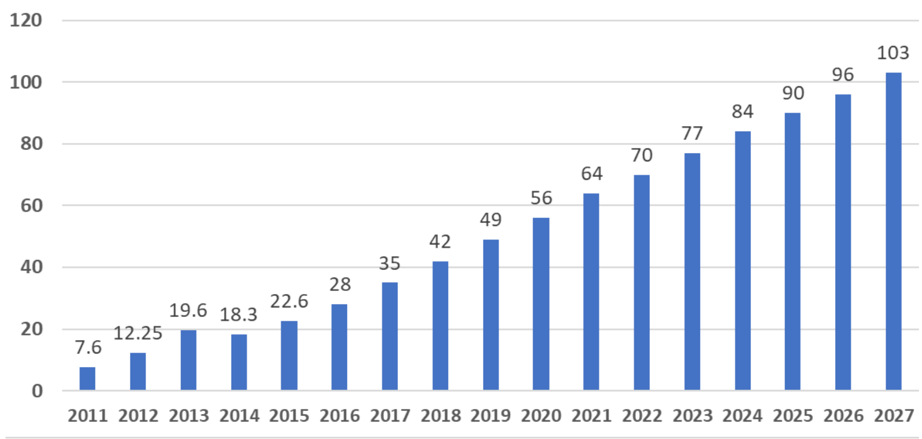


Fig. 3. Big data market size revenue forecast worldwide from 2011 to 2027 (in Billion US Dollars) [32]

Companies that successfully develop their quantitative assessment capabilities can set themselves apart from their competition. More specifically, IBM reported that 53 percent of the industrial manufacturers report that the use of information (including big data) and analytics is creating a competitive advantage for their organizations, compared with 63 percent of cross-industry respondents (survey of 1,144 business and IT professionals in 95 countries, including 124 respondents from the industrial manufacturing industry, or about 11 percent of the global respondent pool). In the industrial manufacturing industries, the percentage of respondents reporting a competitive advantage rose from 33 percent in 2010 to 53 percent in 2012, a 61 percent jump in just two years [31].

The Big Data Market Share Big Data Market Share in Billion US Dollars is presented in the following chart as presented in Fig. 3.

The evolution of AI [33], which can be divided into three generations as shown in Fig. 4, has enabled the development of big data analytics.

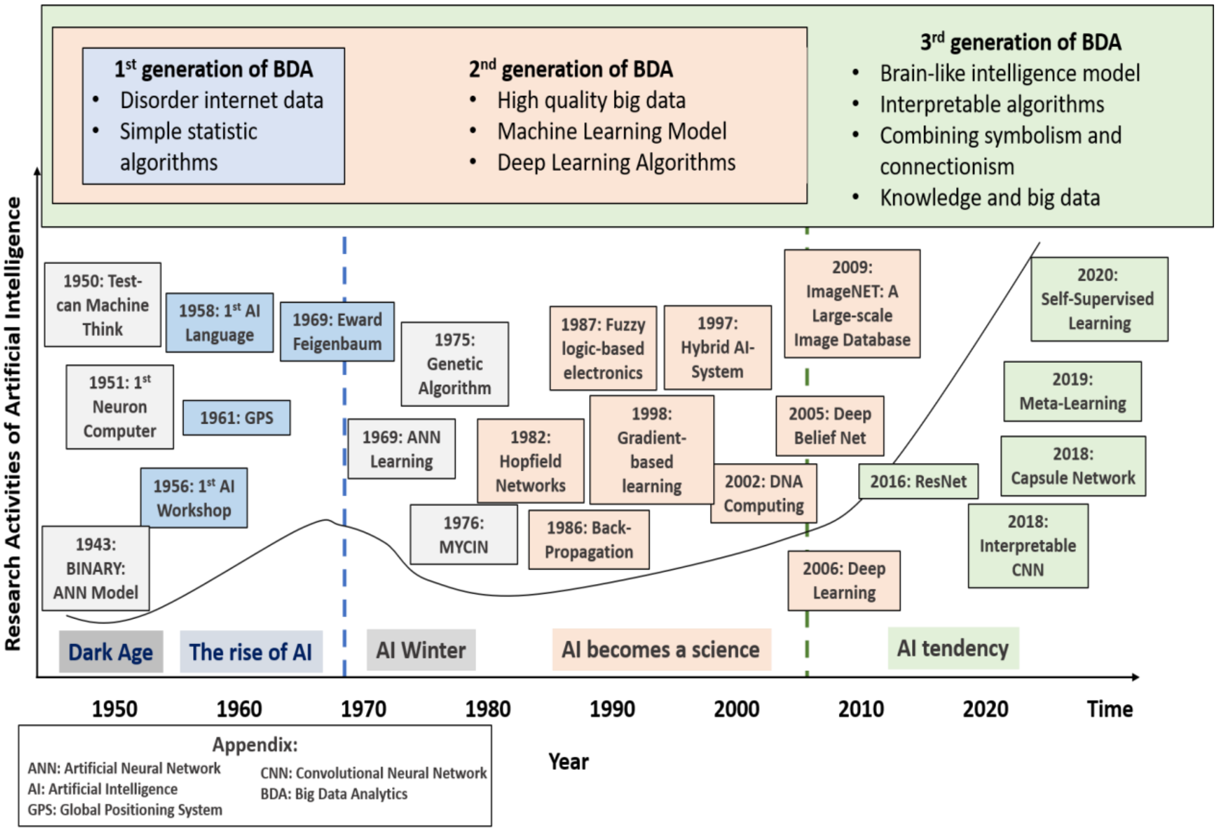


Fig. 4. The development of big data analytics [33]

2.2. INFORMATION TECHNOLOGY AS THE BASIS FOR TRANSFORMATION INTO A DIGITAL SOCIETY AND INDUSTRY 5.0

Current and future industrial and production systems must include interoperability, virtualization, decentralization, real-time capability, and modularity. According to the Boston Consulting Group, these traits are founded on nine pillars as follows [34]:

- Multi-Agent Systems (MAS)
  - Autonomous Robots
  - Artificial Intelligence

- System Integration
- Big Data and Analytics
- Simulation
- Cyber Security
- Cloud Computing
- Additive Manufacturing
- Augmented Reality
- Internet of Things (IoT)

The Smart Factory is a sophisticated system that integrates the main pillar technologies of Industry 4.0 (e.g., autonomous robots, IoT, Big data, Cloud Computing, and simulation). The Smart Factory is based on the idea that traditional centrally managed manufacturing processes will be replaced by decentralized control, in which intelligent machines, robots, tools, and intelligent workpieces continuously communicate and collaborate with one another. Smart Factories are more competitive because they are self-organizing, self-optimizing, and self-organizing. Factories can self-optimize their own performance, self-adapting to new situations and conditions [15].

German industry invested 40 billion euros per year in Industry 4.0 initiatives in 2020. Furthermore, Germany, Switzerland, and Austria (GSA) are investing heavily in digitalization efforts that will enable industrial sectors to capitalize on the opportunities provided by Industry 4.0 technologies. The GSA region's IoT market is worth approximately 36 billion euros. Germany, as the region's largest economy, accounts for 24 billion euros. The Swiss market is worth 7.6 billion euros, while total IoT spending in Austria is currently at 4.2 billion euros. Germany now has 7.6 robots for every thousand industrial workers, which is significantly higher than the European average of 2.7. Additionally, 54 percent of German companies do not use any technology to collect data with the goal of improving production processes.

German firms have also been accused of failing to fully realize the potential of smart sensors and other infrastructure equipment that underpins Industry 4.0. Next, Switzerland is currently the most technologically advanced GSA country. The Swiss government, on the other hand, has been less active in efforts to promote Industry 4.0 and has invested less in digitizing its own services. As a result, electronic interaction between businesses and the public sector in Switzerland is relatively limited. While the government provides funding for Industry 4.0 technologies, it has yet to bring together executives from the public and private sectors to discuss appropriate policy responses. Finally, in Austria, the Ministry of Transport, Innovation, and Technology has encouraged business digitalization. The government also has programs to help small and medium-sized businesses (SMEs) [35].

Currently, two visions for Industry 5.0 are emerging. The first is “human-robot collaboration”. Robots and humans will collaborate whenever and wherever possible in this vision. Humans will concentrate on tasks that require creativity, while robots will handle the rest. The second vision for Industry 5.0 is Bioeconomy [36]. Table 1 compares the visions of Industry 4.0 and Industry 5.0. It is stressed out that Industry 5.0 could refer to both human-robot collaboration and the bioeconomy as well. The application of all the technologies listed in Fig. 5 will help society achieve Industry 5.0 [36].

Table 1. A Comparison of Industry 4.0 and Industry 5.0 Visions [36]

	Industry 4.0	Industry 5.0 (Vision 1)	Industry 5.0 (Vision 2)
Keyword	Smart Manufacturing	Human-Robot Co-working	Bioeconomy
Motivation	Mass Production	Smart Society	Sustainability
Involved Technologies	Internet of Things (IoT) Cloud Computing Big Data Robotics and Artificial Intelligence (AI)	Human-Robot Collaboration Renewable Resources	Sustainable Agricultural Production Bionics Renewable Resources
Involved Research Areas	Organizational Research Process Improvement and Innovation Business Administration	Smart Environments Organizational Research Process Improvement and Innovation Business Administration	Agriculture Biology Waste Prevention Process Improvement and Innovation Business Administration Economy
Power Source	Electrical power Fossil-based fuels Renewable power sources	Electrical power Renewable power sources	Electrical power Renewable power sources

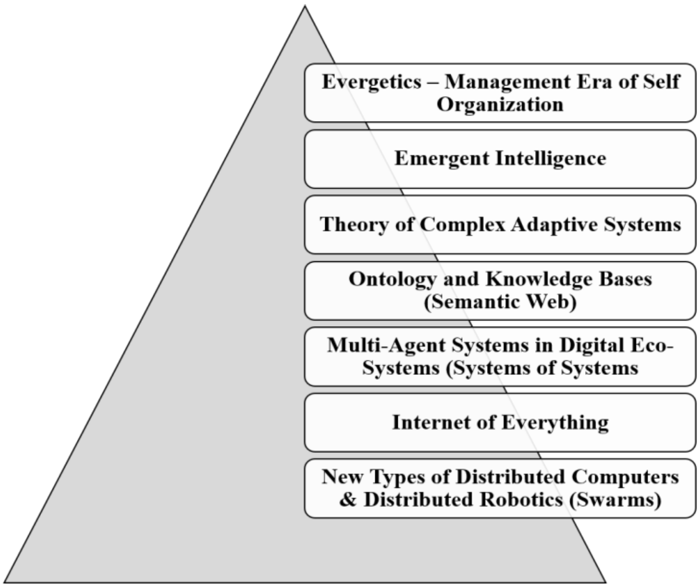


Fig. 5. The convergence of science and technology in Industry 5.0 [36]

New types of distributed Computers & Distributed Robotics

These technologies serve as the hardware foundation for constructing various types of intelligent self-organizing systems. Distributed computer networks enable multi-threading and asynchronous computations as well as the expansion of computing resources.

Internet of things and people

The Internet of Things (IoT), which includes the industrial Internet of Things (IIoT), is a rapidly evolving technology that complements the traditional and well-known Internet and serves as the foundation for automation in Industry 4.0 and Industry 5.0. The Internet of Things (IoT) is a global infrastructure for the information society that enables the interconnection (physical and virtual) of things using existing and emerging information and communication technologies.

In 2013, Cisco coined the phrase “Internet of Everything”. It is thought to be broader than the Internet of Things. The Internet of Everything (IoE) is defined by Cisco as the network connection of people, data, processes, and things [37]. Smart sensors, group robots exchanging wireless signals and working as a single system, RFID tags and so on, are among the promising technologies needed to implement IoT [38]. When Big Data is applied to these technologies, unique solutions emerge. The most important components of Industry 5.0 are likely to be Big Data, IoT, and IoE. During production, all information collected in physical space in the form of Big Data is sent to cyberspace. The difference between Industry 4.0 and 5.0 is the scale at which the latest digital technologies are implemented.

Multi-agent systems and technologies

Multi-agent technologies enable the solution of problems that are difficult to solve using traditional mathematical methods. The introduction of Internet of Things (IoT) technologies assume that everything will be equipped with a small sensor that allows data to be transferred from it to the Internet. Data transfer to the Internet will enable the creation of virtual models of real-world objects, such as factories, whose operations can be monitored in real-time. To communicate between the virtual and real worlds, intelligent agents are used. A multi-agent system is defined as a network of related agents that solve specific problems in a shared environment and interact with one another to achieve the system objectives. In a multiagent system, communication between agents can be accomplished in a variety of ways [39].

Ontology and knowledge bases

There is no universally accepted definition of ontology. Ontology is the formal presentation of knowledge based on conceptualization [40]. Ontology is a set of terms organized into a taxonomy, their definitions, and attributes, as well as the axioms and inference rules that go with them [41]. The ontological approach is successfully applied by multi-agent systems. Ontologies are used to store all an agent's knowledge. They form a knowledge base that includes concepts, subject knowledge, and problem-solving techniques. All of this allows agents to share the information they've gathered and aids in decision-making. At the same time, three qualitatively different methods for integrating the ontology approach can be chosen:

- knowledge and ontologies formed by agents are only accessible to the agents
- knowledge and ontologies formed by agents are combined and stored by one of the agents
- some knowledge and ontologies are stored centrally, while others are distributed among agents

Theory of complex adaptive systems

Aggregation (a hierarchy of elements in which lower-level elements form upper-level elements – aggregates), non-linarites, resource flows, and diversity (lack of equilibrium status) are all characteristics of a comprehensive adaptive system. The theory of complex adaptive systems is the foundation of multi-agent systems. It provides a link between multiagent and nonlinear systems [36]

Emergent intelligence

Emergent intelligence, i.e. intellectual resonance, swarm intelligence, is the manifestation of unexpected properties in a system that lacks any of its individual elements. A key feature is the dynamic and unpredictability of the decision-making process. Multi-agent technologies that implement the interaction of simple elements in their self-organization to solve specific problems are frequently associated with emergent intelligence [42].

Evergetics

Evergetics is an interdisciplinary science that should draw from the humanities and social sciences, as well as control theory, computer science, and other fields. The fact that a person in evergetics is a subject with ways and resources for resolving conflicts and making decisions contributes to its multidisciplinary nature [43]

3. BIG DATA ANALYTICS BASED ON SEMANTICS

The amount of data has grown at an exponential rate in recent years. It is reported that over 16 zettabytes of useful data were produced up to 2020. As a result, the natural expansion of these datasets within organizations necessitates new requirements in terms of processing and exploitation methods, techniques, and tools. The Big Data concept has emerged because of this increase in data. Initially, it was used to describe these massive sets, which typically consist of large amounts of unstructured data that must be stored, processed, and analyzed at high speeds. To that end, the main challenges identified are acquiring, cleaning, integrating, exploiting, analyzing, and visualizing large amounts of data from widely dispersed data sources [44].

The data itself is confronted by Big Data issues, posing challenges at every stage of the value chain, from data collection to visualization and application. As a result, a semantic context is required to assist scientists in gaining access to data, as well as using and interpreting the results. Semantic technologies are used to resolve inconsistencies, evaluate, and discover new information from an existing knowledge base. In the following subsections, discuss the various approaches to combining Semantic with large data in order to connect these data to the real world.

3.1. FROM DATA TO WISDOM: CREATING & MANAGING KNOWLEDGE

The process begins with raw data, which includes signals, symbols, and values that have been collected and observed. The following step is to attempt to convert the raw data into useful information. They can make better decisions with the knowledge they gain from this information (Wisdom). The Data-Information-Knowledge-Wisdom (DIKW) hierarchy or pyramid, connects data, information, knowledge, and wisdom in four layers as illustrated in Fig. 6. The foundation of the pyramid is data, followed by information, knowledge, and wisdom at the top. The characteristics of each data are summarized as follows [45]:

- Data: Are observed, Collected as Facts, Signals and Symbols.
- Information: Meaningful, purposeful, and relevant data. It will answer interrogative questions.
- Knowledge: Provides framed, contextual information, expert insights, and even grounded intuition.
- Wisdom: Adds value, which requires the mental function that we call judgment.

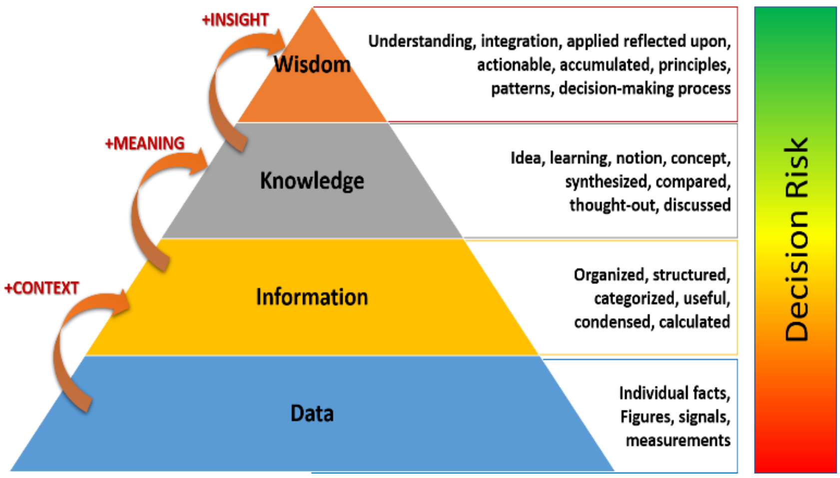


Fig. 6. The data–information–knowledge–wisdom (DIKW) hierarchy as a pyramid to manage knowledge [39]

3.2. TOWARDS WISDOM MANUFACTURING

The Future Internet (FI) is becoming a strategic focus of research in Europe and other parts of the world. According to the European Union's Framework Programme (FP), the future networked society will be supported by four pillars: a) Internet by and for People (IbfP), b) Internet of Contents and Knowledge (IoCK), c) Internet of Things (IoT), and d) Internet of Services (IoS). Without a doubt, the FI will have a significant impact on Computer Integrated Manufacturing (CIM) and Enterprise Integration (EI). In fact, the so-called wisdom manufacturing (WM) [46] has been proposed as an analogy for future networked manufacturing, as shown in Fig. 7.

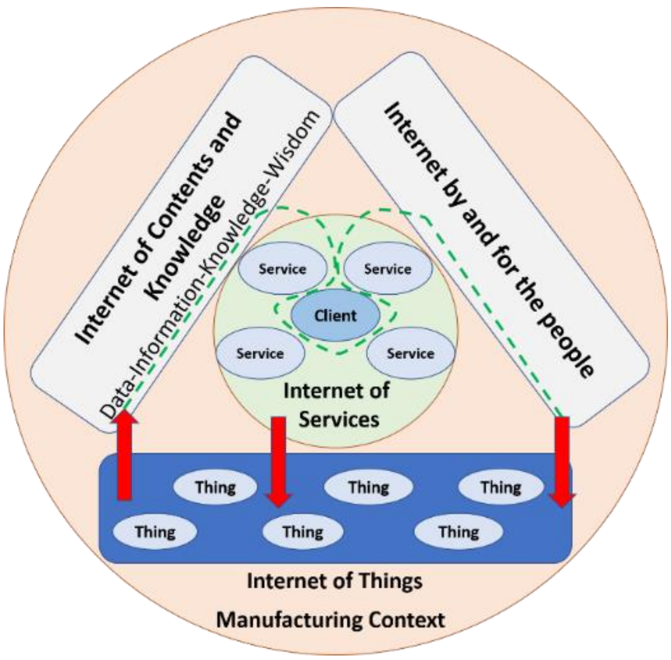


Fig. 7. Wisdom Manufacturing (WM) overview [40]

Moving on, Oxford dictionaries [47] defines wisdom as ‘the quality of having experience, knowledge and good judgement’. Five subcomponents of wisdom most commonly cited in order from high to low are listed as follows:

- decision making/knowledge,
- prosocial attitudes,
- self-reflection,
- acknowledgment of uncertainty,
- emotional homeostasis.

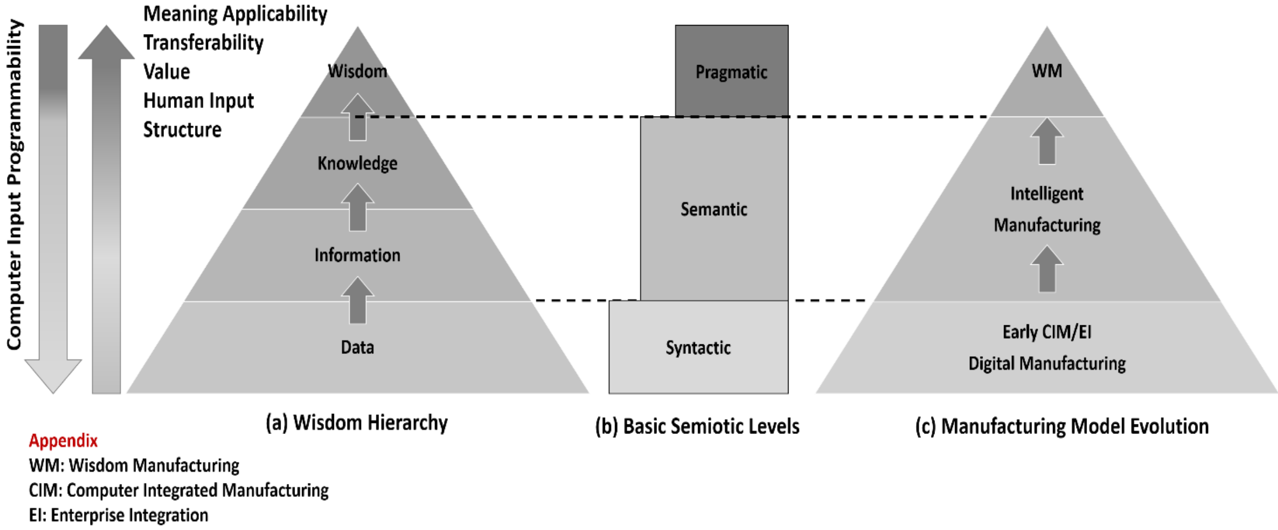


Fig. 8. M turing models corresponding to the DIKW and semiotic levels [48]

Moreover, Fig. 8 presents the “DIKW Pyramid”, the “Knowledge Hierarchy”, the “Information Hierarchy”, and the “Knowledge Pyramid”, and refers to a class of models for representing structural and/or functional relationships between data, information, knowledge, and wisdom, where information is defined in terms of data, knowledge in terms of information, and wisdom in terms of knowledge [48].

3.3. SEMANTICS FOR BIG DATA ACQUISITION

The semantic aspect can be added to the acquisition step, which represents the process of gathering, filtering, and cleaning data before storing it. Furthermore, data analysis can be triggered to capture relevant and valuable information. As a result, effective analytical algorithms are required to comprehend the data source, continuously process the data, and reduce the data prior to storage [49]. The authors in [50] proposed a scheme for retrieving and storing large amounts of Resource Description Framework (RDF) data in Hadoop Distributed File System (HDFS). The system parses RDF data and saves it to a file. The authors then propose an algorithm for determining the best query plan for executing multiple queries on datasets of various sizes. As a result, this system is scalable and efficient enough to handle billions of RDF triples with ease. The authors in [51] include three layers of MDA (Model Driven Architectures) in their paper: the CIM (computational dependent model), the PIM (platform independent model), and the PSM (platform specific model) (platform specific model). These layers are then converted into a Big Data architecture, which entails identifying a set of predefined parameters. The PIM and PSM are then mapped to these models (depending on the accurate deployment constraint). Finally, the TOREADOR platform can be used to run them. As presented in Fig. 9, the system optimizes reuse, reduces development costs, and provides a programming interface to interact with distributed memory (using the spark framework, which implements a fault-tolerant distribution in a cluster).

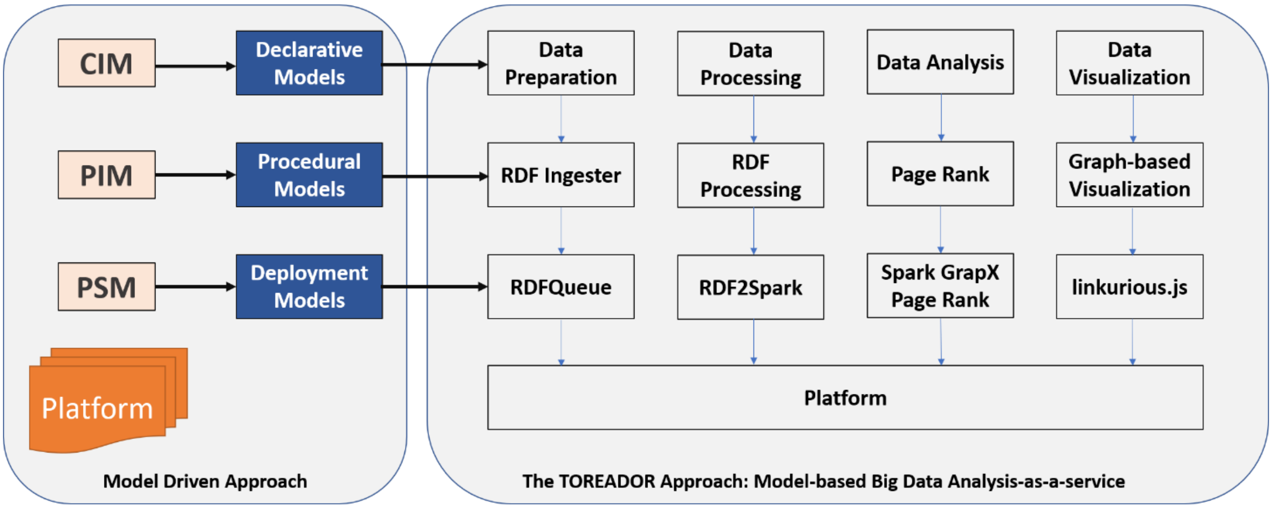


Fig. 9. Model Driven Approach [51]

3.4. SEMANTICS FOR BIG DATA INTEGRATION

It relates with the semantic dimension being introduced after the data acquisition and cleansing step is done in the integration phase. It has to do with the clean integration and aggregation of unstructured Big Data, which lacks convincing information. These data integration techniques aim to combine data from various sources and provide users with a unified view of the data being analysed. The authors in [52] propose a method for dealing with large amounts of semantic data. It has four layers: the metadata layer, which describes tourism data such as geographic location and relationships between resources using different metadata rules such as MARC, DC, and GILS. The ontology layer ensures semantic interoperability across a variety of metadata types. To this end, two options are available: the first is to use the OWL knowledge representation language to integrate the attributes and concepts of the various metadata rules into an ontology. The second method involves converting the metadata format to RDF using an ontological language. The data is published using the Linked Data principle by the Linked Data layer. It creates semantic interoperability between data by providing a unified access mechanism for various data formats. The data application layer enables Interactive research to provide both traditional keyword-based retrieval and a more user-friendly interface. Next, using semantic integration techniques, a system for combining different sources of air traffic management data is presented. It converts data from various sources into a standardized semantic illustration in a triples store using ontology-based sherlock data. Another approach uses the Semantic Sensor Network ontology to transform sensor data into semantic data [53]. In this study, the authors proposed several methods for storing and processing semantic data. The semantic data obtained is analysed to uncover useful information. Furthermore, Big Data technologies are used to transform raw sensor data into semantic data and achieve the required scalability in intensive data scenarios. The application of the SSN (Semantic Sensor Network) ontology to sensor data adds semantic compatibility. The primary issue is the increasing complexity of sensors used in new systems.

Large data sets, by definition, contain a lot of unstructured information. As a result, large-scale processing of such semi-structured or unstructured datasets poses a significant challenge [54]. The authors in [55] propose an empirical investigation of character-convolutive networks for text classification. ConVnets were created by the authors using an English thesaurus obtained from WordNet, in which each synonym for a sentence or a word is classified by semantic proximity in the most widely understood sense. Next, the authors in [50] propose a method for selecting an appropriate model for analysing Big Data using semantic technologies. In order to aid inference for semi-automated model selection, an analytics ontology is created using the Ontology Web Language (OWL). In the analysis phase, a hierarchical workflow for predictive analytics is created. The first step in the workflow is to choose a practical model based on the dataset properties. In the second step, the data is prepared for analysis. The third step is to fine-tune the models that will be used. A set of predictor variables is suggested for each step. Next, an ontology-based workflow generation approach is proposed in [56] to automatically generate the workflow depicted in Fig. 10.

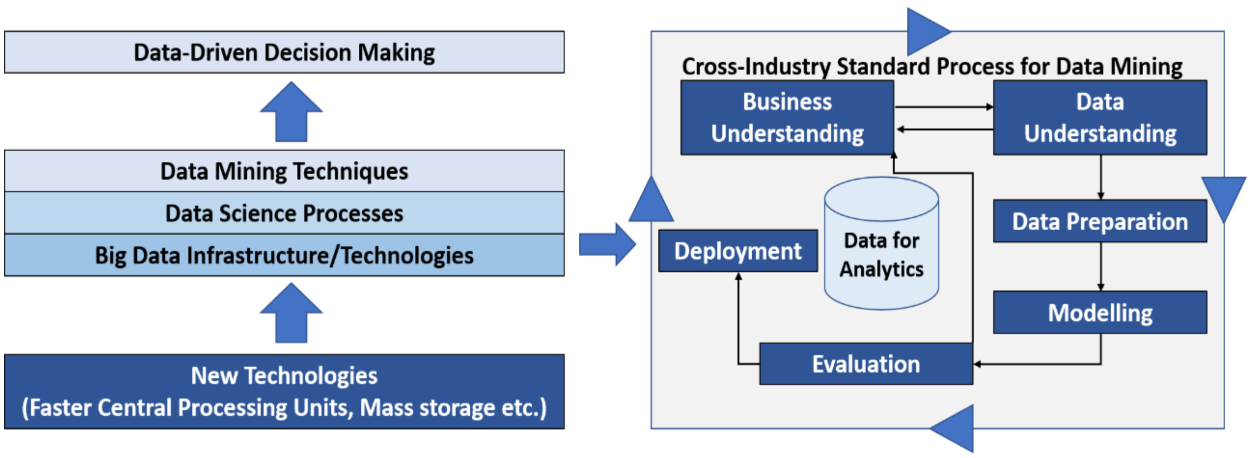


Fig. 10. Process for Big Data Analytics [51]

The Cross-Industry Standard Process for Data Mining (CRISPDM) method is computerized using the Automatic Service Composition (ASC) method. Two different workflow approaches are developed based on an ontology designed for CRISP-DM. To begin, a rule-based approach to inferences is proposed. It uses Semantic Web Rule Language (SWRL) rules to find abstract services based on the properties of datasets and the requirements of the users. Second, two use case scenarios are used to demonstrate the feasibility of a query-based approach. As a result, this approach can be used to generate more application-specific abstract workflows. Moving on, the authors in [58] propose a novel framework for combining semantic methods and Big Data processing in Big Data security analysis. To begin, the framework gathers and preprocesses a large amount of data from various sources. Then, it provides large spaces to store the collected data. The data is then processed using a variety of tools and libraries. Finally, the framework describes three types of analysis: data analysis (data mining, machine learning, and statistic analysis methods), and statistical analysis methods. Human Computer Interaction (HCI) analysis, which employs semantic analysis and data analysis, and semantic analysis, which is based on the use of ontology. This framework improves security analysis techniques such as real-time Big Data, computational performance, batch processing, data association, and data mining in order to meet data volume performance requirements. In order to address these gaps, the authors in [45], propose a method that employs semantic memory to aid in the semantic classification of data involving the value and type characteristics as presented in Fig. 11.

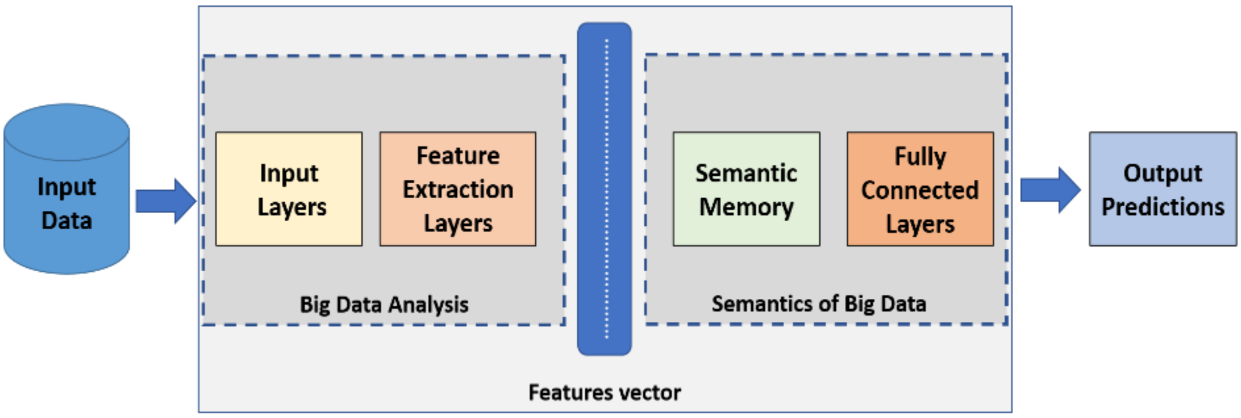


Fig. 11. Semantic Memory for Big Data Analysis Architecture [45]

3.5. AUTOMATED PREDICTIVE BIG DATA ANALYTICS USING ONTOLOGY BASED SEMANTICS

Predictive big data analytics is based on decades of statistical and machine learning advancements. Several frameworks are being developed to support large-scale data analytics. Drill, Hadoop, Mahout, Storm, Spark, and SCALATION are among the members of the group. These frameworks aim to speed up computation and support a larger volume of data by utilizing databases and distributed file systems, as well as parallel and distributed processing. Many complex steps are involved in predictive big data analytics, many of which necessitate a high level of expertise. The hierarchical workflow for the predictive analytics process to help manage the complexity is presented in Fig. 12 and depicts the top level of our hierarchical workflow.

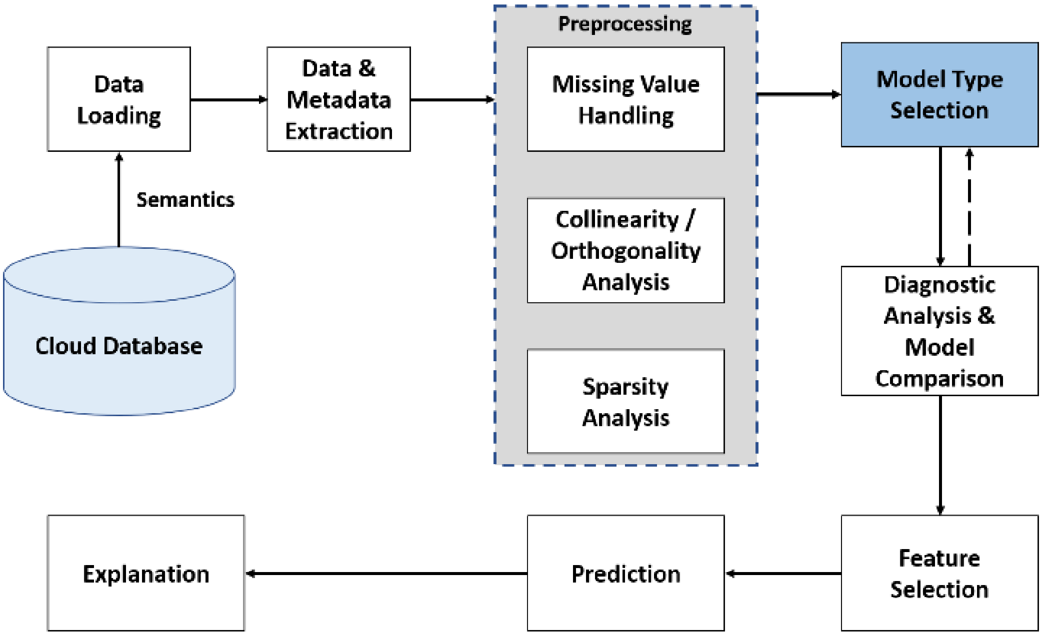


Fig. 12. Predictive Analytics Workflow [59]

4. ARTIFICIAL INTELLIGENCE APPROACHES IN BIG DATA ANALYTICS

In the following paragraphs, an analysis of AI technologies in Big Data analytics is performed. Ultimately, the purpose of this Section, is to identify the correlation between AI, Big Data analytics and most importantly Semantics, in an attempt to unify the information context of manufacturing processes, aiming to provide a framework for process optimization. By extension, the combination of AI technologies and techniques is already facilitating engineers to gain useful insights for various aspects of manufacturing and production systems, with the installation of suitable monitoring and sensing systems (e.g., Wireless Sensor Networks – WSN). However, in order to transform raw data into knowledge, and ultimately to wisdom, the combination of Ontologies with AI technologies will play a key role.

4.1. APPLICATIONS OF DEEP LEARNING IN BIG DATA ANALYTICS

The raw data in Big Data systems is becoming increasingly diverse and complex, consisting of un-categorized/unsupervised data along with a small amount of categorized/supervised data. Working with the diversity of data representations in a repository presents unique challenges for Big Data, which necessitates Big Data pre-processing of unstructured data in order to extract structured/ordered data for human and/or downstream consumption. Beyond the four Vs, Big Data Analytics faces a number of challenges. Some key problem areas include: data quality and validation, data cleansing, feature engineering, high-dimensionality and data reduction, data representations and distributed data sources, data sampling, scalability of algorithms, data visualization, parallel and distributed data processing, real-time analysis and decision making, crowdsourcing and semantic input for improved data analysis, tracing and analyzing data provenance, data discovery and integration, parallel and distributed computing, exploratory data analysis and interpretation, integrating heterogenous data, and developing new models for massive data computation [60]. Deep Learning algorithms extract meaningful abstract representations of raw data using a hierarchical multi-level learning approach, in which more abstract and complex representations are learned at a higher level based on less abstract concepts and representations learned at lower levels of the learning hierarchy. While Deep Learning can be used to learn from labelled data, if there is sufficient amount of it, it is best suited for learning from large amounts of unlabelled/unsupervised data, making it ideal for extracting meaningful patterns and representations from Big Data [61]. Once Deep Learning has learned hierarchical data abstractions from unsupervised data, more traditional discriminative models can be trained with fewer supervised/labelled data points, where the labelled data is typically obtained through human/expert input. Deep Learning algorithms have been shown to perform better than shallow learning architectures at extracting non-local and global relationships and patterns in data [62]. Other useful characteristics of Deep Learning include: (1) relatively simple linear models can effectively work with knowledge derived from more complex and abstract data representations, (2) Increased automation of data representation extraction from unsupervised data allows it to be applied to a wide range of data types, including image, textural, audio, and so on, and (3) relational and semantic knowledge can be obtained at higher levels of abstraction and representation of the raw data. Deep Learning algorithms and architectures are better suited to address issues related to Volume and Variety of Big Data Analytics when considering each of the four Vs of Big Data characteristics, i.e., Volume, Variety, Velocity, and Veracity. Where algorithms with shallow learning hierarchies fail to explore and understand the higher complexities of data patterns, Deep Learning inherently exploits the availability of massive amounts of data, i.e. Volume in Big Data. Furthermore, because Deep Learning deals with data abstraction and representations, it is likely suited for analysing raw data presented in various formats and/or from various sources, i.e. Big Data variety, and may reduce the need for human experts to extract features from each new data type observed in Big Data [63].

4.2. SEMANTIC INDEXING

Information retrieval is a key task associated with Big Data Analytics (BDA). Large-scale quantities of data such as text, image, video, and audio are being collected and made available across various domains, e.g., social networks, machine monitoring and so on, making efficient storage and information retrieval a growing problem in Big Data. Massive amounts of data are available in these systems, which require semantic indexing rather than being stored as data bit strings. Semantic indexing improves the efficiency of data presentation and makes it more useful as a source for knowledge discovery and comprehension, for example, by making search engines work faster and more efficiently.

Deep Learning can be used to generate high-level abstract data representations that will be used for semantic indexing instead of using raw input for data indexing. These representations can reveal complex associations and factors, leading to semantic knowledge and understanding (especially when the raw input is Big Data). However, it should be noted that in order to actually grant a good semantic understanding and comprehension of the input, the high-level abstract data representations must be meaningful and demonstrate relational and semantic association. While Deep Learning aids in the semantic and relational understanding of data, a vector representation (corresponding to the extracted representations) of data instances would allow for faster searching and retrieval of information. For many domains, document (or textual) representation is an important aspect of information retrieval. The goal of the document representation is to create a representation that condenses specific and unique aspects of the document, such as the document topic. Word counts, which represent the number of times each word appears in the document, are used in document retrieval and classification systems. In such document representation schemas, individual words are considered dimensions, with different dimensions being independent.

Hinton et al. [64] present a Deep Learning generative model for learning document binary codes. The lowest layer of the Deep Learning network represents the wordcount vector of the document, which is high-dimensional data, and the highest layer represents the document's learnt binary code. Further to that, Deep Learning generative models can also be used for the generation of shorter binary codes by limiting the number of the variables used by the highest layer in the learning hierarchy. The memory addresses for these shorter binary codes can then be used. Additionally, semantic hashing techniques are appealing for information retrieval because documents that are similar to the query document can be retrieved by locating all memory addresses that differ by a few bits from the query document's memory address. Next, the authors in [65] describe a study in which Deep Learning model parameters are learned using both supervised and unsupervised data. The benefits of this strategy are that there is no need to completely label a large collection of data (because some unlabelled data is expected) and that the model has some prior knowledge (via the supervised data) to capture relevant class/label information in the data. Another technique for automated extraction of semantic representations from Big Data is “word2vec” tool by Google. The word vectors are generated from a large-scale text corpus using this tool. It builds a vocabulary from the training text data and then learns vector representations of words, which can then be used as features in many Natural Language Processing (NLP) and Machine Learning Applications [66].

4. 3. BIG DATA OPTIMIZATION

The growing popularity of Big Data applications, many of which require the processing of large amounts of data arriving at high speeds from streaming data sources, has opened up new opportunities for dynamic multi-objective optimization. The content of Big Data has been increasing in recent years, and the goal of Big Data analytics has also changed over time. More specifically, not only should the algorithm be able to handle dynamically changing data, but it should also be able to adjust the data analytic target. One of the most important characteristics of Big Data is that the data was collected from various sources in order to create a large dataset. In most cases, multiple objectives must be met at the same time in these large datasets. Traditional methods can only be used on continuous and differentiable functions, and they require a series of separate runs to meet various goals [60]. Several approaches based on adapting metaheuristic techniques to work in parallel on Hadoop ecosystems have been proposed in the last decade. These suggestions are for data mining or data management applications, such as:

- a swarm intelligence method to optimize the feature selection in Big gene expression datasets [68],
- data partitioning in Big Databases [69],
- dimension reduction in Big Data analytics [70],
- pattern detection with Artificial Immune Algorithms [71],
- a parallel MapReduce evolutionary algorithm for graph inference [72],
- and a parallel artificial ant colony optimization for task scheduling in clusters environments [73].

Therefore, the four components of Big Data analytics and specifically Big Data optimization are as follows: a) handling large amounts of data, b) high dimensional data, c) dynamical data, and d) multi-objective optimization. The majority of real-world Big Data issues can be modelled as large-scale, dynamical, and multi-objective issues. This sub-section presents in a brief manner the principles of optimization using dynamic multi-objective algorithms, metaheuristics, for dealing with Big Data optimization problems [74].

The optimization in Big Data as one of the most important opportunities and challenges in the field [75]. Big Data optimization is a novel research area in which traditional problems, and thus traditional metaheuristics, must account for the newly created amount of data in their context. For example, a few years ago, The Travelling Salesman Problem (TSP) would only be created using static data from a city and would not adapt to real-time changes in the city (traffic jam, work on roads and so on). However, thanks to Open Data websites, we can now change the data of the problem in real-time via streaming, resulting in better and more realistic routes. Because of the nature of the problem, optimization algorithms for resolving Big Data optimization should be able to manage streaming data sources, and the algorithms should be able to solve multi-objective problems because more than one objective must be satisfied at the same time. As a result, they have to be Dynamic Multi-Objective algorithms.

4.4. SEMANTIC WEB

Semantic Web is an activity led by World Wide Web Consortium (W3C) to enhance documents available on WWW so that they would have meaning understandable to computers. The languages such as Resource Description Framework (RDF) and Web Ontology Language (OWL) allow us to describe the data in a way that they can be easily integrated and queried. The key effort is to use semantic web technology (via ontology) to represent all Big Data Analytics (BDA) knowledge. The next step is to use reasoners or SPARQL queries to fetch this information or derive new knowledge from it. The main technologies used in Semantic Web models related to BDA are presented in brief as follows: a) Ontology, b) Resource Description Framework (RDF), c) Ontology Web Language (OWL), d) SPARQL and e) Semantic Web Rule Language (SWRL).

The main idea is to use semantic web technology (via ontology) to represent all Big Data analytic knowledge, and then use reasoners or SPARQL queries to fetch this information or derive new knowledge from it. The main technologies used in Semantic Web models related to BDA are presented in brief as follows:

- **Ontology.** As per the definition in [76], a formal representation of the real world is provided by an ontology. Ontologies define data models in terms of classes, subclasses, and properties, which define an explicit description of concepts in a domain of discourse (classes or concepts), properties of each concept describing various features and attributes of the concept (properties), and restrictions on properties. Ontologies are part W3C (World Wide Web Consortium) standard stack of the Semantic Web. A knowledge base is made up of an ontology and a set of individual instances of classes that provide services to enable interoperability across multiple, heterogeneous systems and databases.
- **RDF.** A W3C recommendation called Resource Description Framework defines a language for describing web resources. RDF statements are used to describe resources in terms of their properties and value. The RDF Schema (RDFS) defines the RDF vocabularies [77].
- **OWL.** The Ontology Web Language, which extends RDF and RDFS by adding a vocabulary, is used to define ontologies on the Web. OWL is equivalent to a very expressive description logic (DL) from a formal standpoint, where an ontology corresponds to a Tbox (terminological component). OWL-DL is a syntactic description that allows for maximum expressiveness while maintaining computational completeness and decidability [78].
- **SPARQL.** It is a query language that allows to quickly access RDF data. SPARQL is a graph-matching query language for extracting knowledge from models [79].
- **SWRL.** The Semantic Web Rule Language adds procedural knowledge to OWL-based ontologies, overcoming some of the limitations of ontology inference, particularly in identifying semantic relationships between people. To represent semantic rules, SWRL employs the common logic expression "Antecedent => Consequent" [80].

The W3C Semantic Sensor Network (SSN) is an Ontology expressed in OWL2 and describes sensors capabilities, measurements, observations, and deployments. Moreover, the key concepts and relations, split by conceptual modules are presented in Fig. 13. IoT adoption in manufacturing allows for the transformation of traditional industrial systems to new digitalized ones, resulting in enormous economic prospects through industry reshaping. Industrial IoT enables modern businesses to more readily implement new data-driven strategies and deal with global competitive pressure.

However, as IoT becomes more widely adopted, the total volume of created data grows, changing industrial data into industrial Big Data. To demonstrate how industrial data may be generated and how it can lead to Industrial Big Data, the authors in [82] developed IoT application is used in a mould-making industry with 100 machine tools and 150 personnel. The designed and developed tool was implemented in one machine tool, and the data captured and communicated to the gateway were measured to determine the volume of data generated.

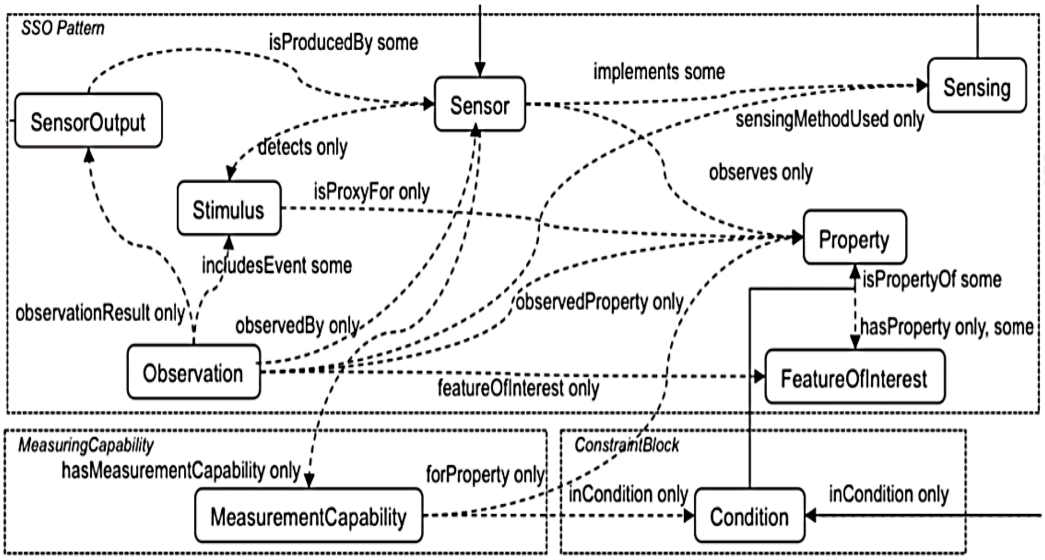


Fig. 13. Semantic Sensor Network (SSN) Ontology structure [81]

The camera is thought to be for ten of the most important machine tools in the manufacturing process. Afterwards, the volume of the generated data per machine tool was calculated, along with the total volume of generated data from the entire shopfloor. The volumes calculated clearly indicate the enormous data volumes that can be generated in a shopfloor using the aforementioned types of sensors. This yields a volume of generated data, which are displayed in (Table 2), in correlation with the sampling rate of each sensor used.

Table 2. Data generated by the developed DAQ [82]

DAQ Level –Generated data		
Sensor	Sampling rate	Megabytes per hour
Spindle Closed-loop Hall sensor	1 MHz	13,733
Axis X split-core CT	1 kHz	13,73
Axis Y split-core CT	1 kHz	13,73
Axis Z split-core CT	1 kHz	13,73
Mains current split-core CT	1 kHz	13,73
Mains voltage insulation Transformer	1 kHz	13,73
Camera	10 screens/min	293
SUM		14,095

The firmware of the proposed monitoring tool is designed to provide and transmit processed data to the Cloud at a rate of 4 measurements per second. Based on this rate, the Table 4, below, presents the meaningful data that resulted from the Data Acquisition (DAQ) processing. These data are generated using the various types of sensors comprising the sensing system of the IoT tool, engineering data, and business data generated in the case of the mould–making industry. The selected transmission rate of the generated data in this case study is based on the production of the mould-making industry.

Table 3. Industrial Data Transmitted [82]

Data source	Generated Data/day	Generated Data/month	Generated Data/year
Machine tool	1,356MB	39.73 GB	0.47TB
Camera	7,200MB	211GB	2.47TB
Shop-floor	204GB	6TB	72TB
Production Network (50 industries)	10TB	300TB	3.6EB

The engineering data contain all the necessary documentation for the products from a manufacturing standpoint. Furthermore, business data includes information from Information Technology (IT) tools (Enterprise Resource Planning, accounting software, and communication software). If we consider data from a variety of IoT tools, data from human operators' mobile devices, and data from various and different IT tools in the company that can be interfaced through the OPC-UA architecture, a large volume and variety of data will be considered, thereby reaching the starting point for Industrial Big Data. One of the most difficult aspects of Industrial Big Data is gathering and considering only the most important data for each decision that must be made. Data can be processed locally by the machine tools nodes and the microcomputer using the proposed monitoring system, reducing the volume of data that must be transmitted and stored in the Cloud database. Moving in this direction, the actual data transmitted will be of low volume, allowing for quick and efficient decision-making.

4 .5. FRAMEWORKS OF BIG DATA ANALYTICS FOR INTELLIGENT MANUFACTURING SYSTEMS

The big data-driven operation of manufacturing systems shifts to a framework of “correlation & prediction & regulation” under the data science paradigm as presented in Fig. 14. From the data point of view, correlation analysis means quantifying the relationship between various factors in manufacturing systems. (2) Prediction refers to using machine learning methods to forecast the performance indicators of manufacturing systems (e.g., cycle time). (3) Regulation is the process of improving the performance of the system by optimizing the controllable variables.

The framework below follows a four-step process in terms of data processing: (1) Manufacturing data is combined and pre-processed to produce data that is both reliable and reusable, (2) Correlative analysis is used to determine the explanatory factors for manufacturing system performance indicators. The fluctuation of system performance indicators can be modelled using the explanatory factors, (3) The performance indicators of the system can be predicted using the explanatory factors to provide information for decision-making. For accurate prediction, various types of machine learning models are developed. Prediction is broadened in this section to include fault detection, classification, and other advanced prediction tasks in manufacturing systems, (4) Decision-making methods can be implemented to improve the performance of the system based on the predicted value.

Design data analysis can usually be used to improve the function, structure, and process of a product. Manufacturing systems can be made more efficient by using planning and scheduling. Process and quality control systems were used to control and improve product yield.

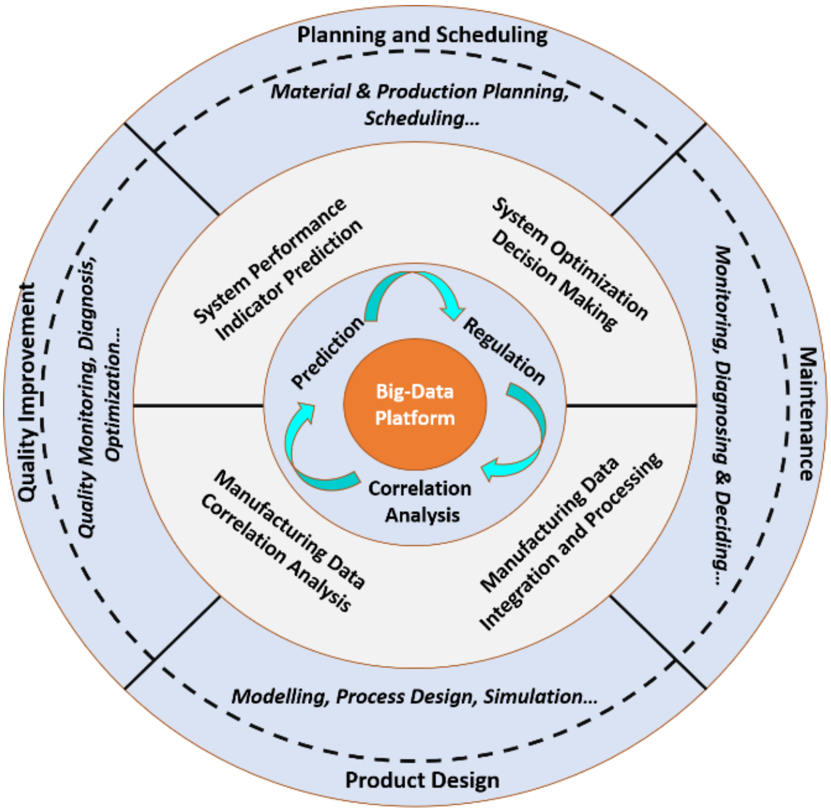
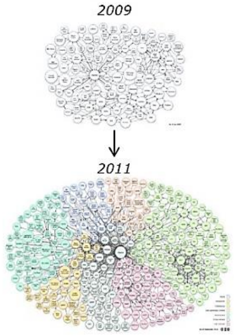
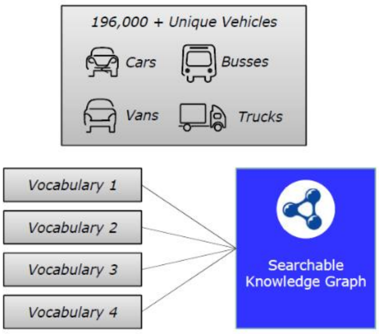
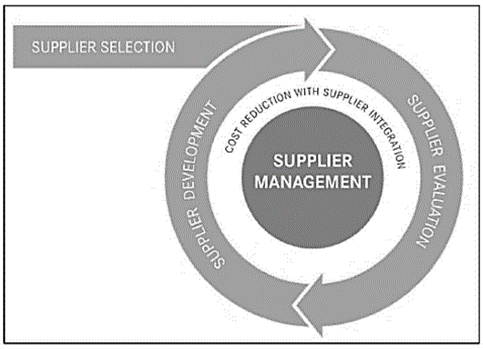
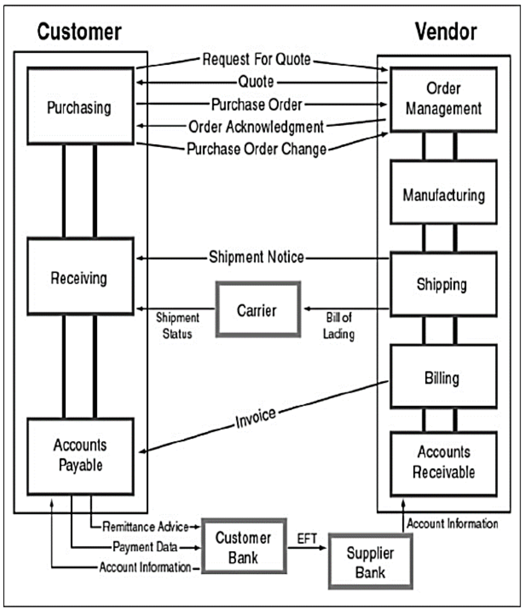


Fig. 14. The framework of big data driven intelligent manufacturing [75]

4. 6. APPLICATION AREAS OF SEMANTICS IN INDUSTRIAL COMPANIES

Semantic Intelligence is a topic that brings together the activities of the AI, Machine Learning, and Semantic Web communities. The selection of an appropriate processing model and analytical technique is a complex TASK that is influenced by the business concerns of the targeted domain, such as sensing and cognition in production plants, automated response in control rooms etc. Advanced analytical services and semantic data lakes integration is a complex research area. The goal of semantic intelligence is to make business intelligence solutions accessible and understandable to humans, rather than making data and processes understandable to machines. Natural language processing (NLP) and semantic analysis, for example, are used in human-machine interactions to interpret and respond to posed inquiries while incorporating semantic information (digital assistants). In this scenario, NLP techniques integrate statistical and linguistic techniques with graph-based AI. The following Table 4 summarizes some key application areas of semantics in industrial companies.

Table 4. Industrial applications of semantics

Industry 4.0	Frameworks / Developments	References / Source
1. Marketing and Sales Application Product catalogues • semantic product catalogues in Open Linked Data • Current topic in Productive 4.0 (EU funded) and Smart Stage (BMWi funded)		[84]
2. Internal Application • Knowledge management, e.g. for large IT projects with many stakeholders • Data warehousing, e.g. for spare parts in a car repair information service		[85]
3. Procurement and Supply Chain Management Appli-cation • Unifying supplier catalo-gues • Knowledge management on potential suppliers and their offers		[86]
4. Business Transaction Application • Information exchange for tender processes / RFQs • Order and Call-off processing		[87]

5. BIG DATA ANALYTICS AND INDUSTRY 5.0 CONCEPTUAL FRAMEWORK

To meet the future manufacturing complexity of increasing customization through an optimized manufacturing (robotized) process, Industry 5.0 recognizes that man and machine must be interconnected. Industry 5.0 is likely to affect the economy, ecology, and the social world. Reduced waste material could have a significant impact on both the economy and the environment as manufacturers move toward zero-waste production, lowering material costs and waste management costs. In a nutshell, Industry 5.0 will completely delegate mechanical tasks to robots, while humans will be in charge of the creative side. The nine pillar technologies of Industry 5.0 are listed below:

- Big Data Analytics
- Augmented Reality
  - Improves Decision Making and Work process
  - Time Saver
  - Cost effective
  - Effort Precision
  - Provides real-time information to workers
- Autonomous Robots/Drones
  - Assist in Manufacturing
  - Usage shall Impact on the Cost Benefit
  - Facilitate in Quality
- Timeline on Delivery
- Simulation
  - Simulation Augments the Precision of Quality,
  - Cost and Process
  - Optimizing the Productivity in Virtual mode
  - Synchronising real-time Data
  - Virtual model
  - Interacting Machines, Work force and Products
- Horizontal and Vertical System Integration
  - Cohesive System Integration
  - Enterprise Data Integration
  - Value Chain Integration
- The Industrial Internet of Things
  - Centralized Controllers with Embedded Computing
  - Analytics and Decision-Making Systems Enabling Real-time Responses
- Cyber Security
  - Deployment
  - Use of Standard Communication Protocols
  - Well Defined Industrial Systems
  - Sophisticated Identity and Access Management
- The Cloud
  - Datta Driven Services for Production Systems

  - Business Process and Workflow on Cloud
  - Data Information and Interchange Across Cloud
- Additive Manufacturing
  - Optimizes workflow
  - Process through 3D Printing on Prototype and Individual Component Production.

Big data-driven intelligent manufacturing applications have begun to emerge, fuelled by intelligent sensing, the Internet of Things (IoT), distributed storage computing, machine learning, and other Industry 4.0 technologies. Applications in product design, planning and scheduling, quality optimization, and equipment operation and maintenance as shown in the framework of Fig. 15.

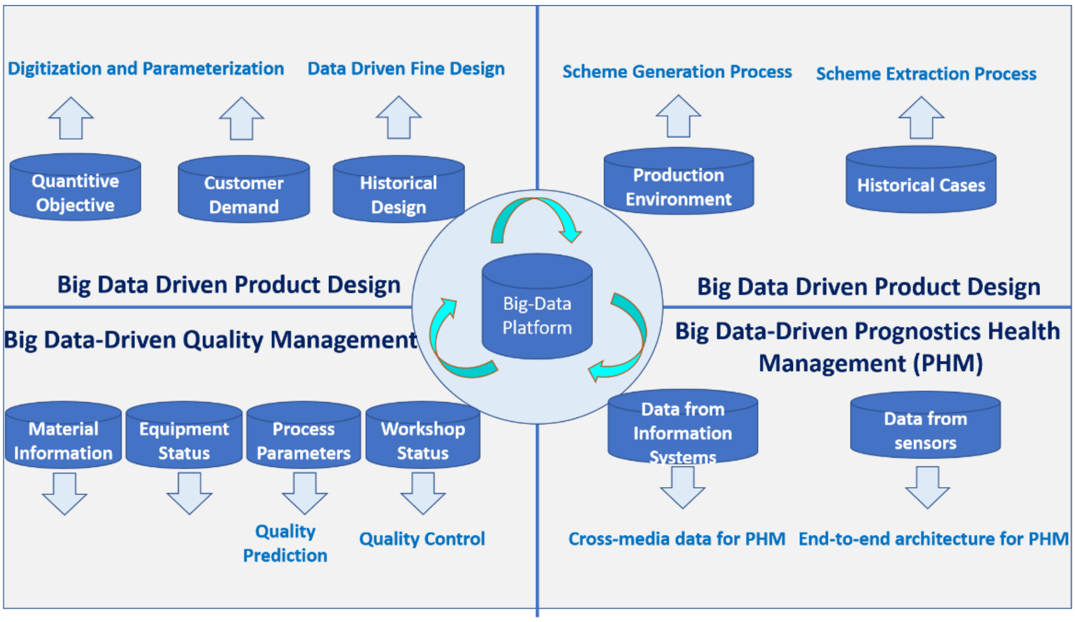


Fig. 15. The applications of BDA in manufacturing systems [88]

6. DISCUSSION

The third and fourth generation (3G, 4G) mobile networks, as well as other communication technologies, are unable to meet the demands of Cyber Physical Manufacturing Systems (CPMS) [77], such as high data rate, high reliability, high coverage, low latency, and so on, obstructing the development and implementation of the system. Manufacturing is the second-most attacked industry, yet the manufacturing sector lags when it comes to security. Smart factories can be subject to the same vulnerability exploitation, malware, denial of service (DoS), device hacking, and other common attack methods that other networks face. And the smart factory's expanded attack surface makes it extra difficult for manufacturers to detect and defend against cyberattacks. These threats now work on an entirely new level with the dawn of the IoT, and they can result in serious physical consequences, especially in the realm of the Industrial IoT. A few of the new security challenges that businesses face in the age of Industry 4.0 are listed hereinafter [89]:

- Every connected device is a potential vulnerability,
- Manufacturing systems, such as Industrial Control Systems (ICS), have specific vulnerabilities that make them especially vulnerable to cyberattacks,
- Because Industry 4.0 connects previously isolated systems, the attack surface expands,
- Upgrades are often installed piecemeal since the systems are very complex,
- Manufacturing has many fewer regulated compliance standards than other sectors,
- Visibility is poor across separate systems and isolated environments.

Therefore, protecting against evolving threats is an active challenge, which is evolving to a full-time job as more connected systems are deployed and the opportunities for an attack on intellectual property increase. As a result, the manufacturing sector needs to [90]:

- Adopt a risk-based security mindset,
- Keep an accurate inventory of all Operational Technology (OT) assets in real-time, Combine the best of IT and OT as an integrated defense strategy across all attack surfaces,
- Identify and fix outdated systems, unpatched vulnerabilities, and poorly secured files,
- Adopt a security-first strategy to the deployment of new connected systems,
- Spot potential threats with real-time vulnerability assessments and risk-based prioritizations,
- Ensure that technology suppliers and connected equipment manufacturers commit to regular security and software patches and audits.

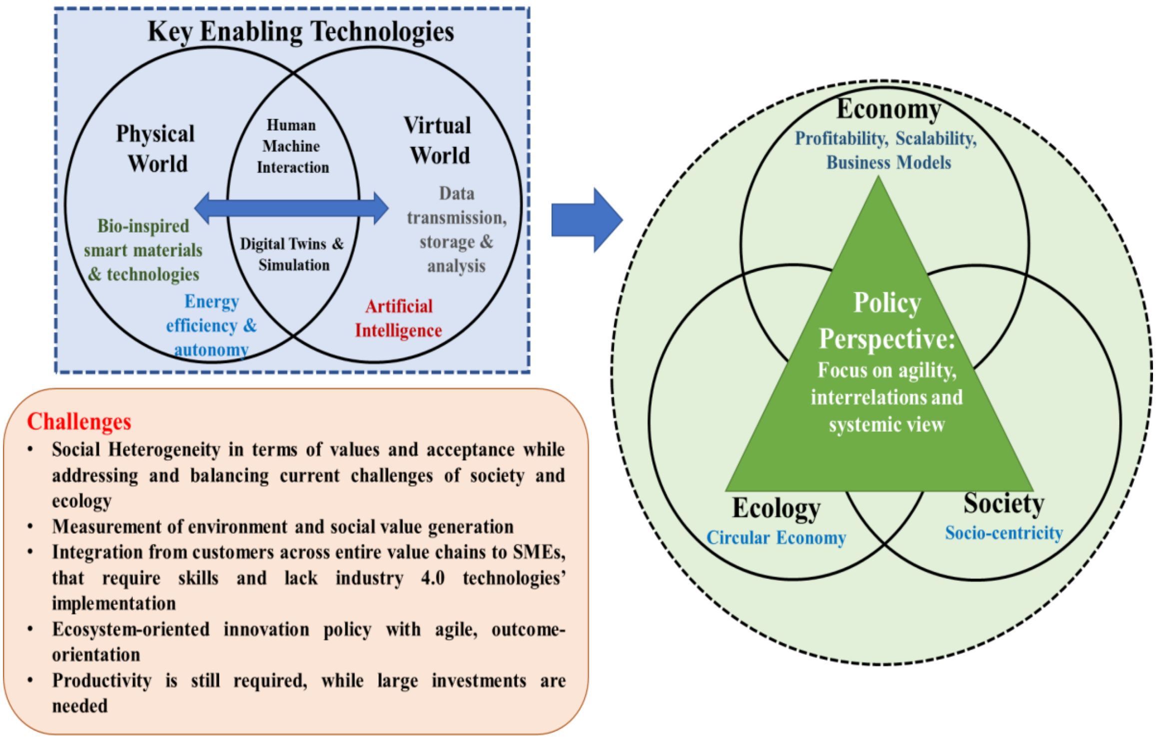


Fig. 16. Key features, the goals, the technological enablers, and the challenges of Industry 5.0 [91]

Similarly, 5G has a lot of potential as a future advanced wireless transmission technology for Industrial Internet of Things (IIoT) and CPMS. Moreover, Industry 5.0 has arrived, and involves combining increasingly powerful and precise machinery with the unique creative potential of the human operator. Industry 5.0 has already triggered unstoppable changes. By combining the capabilities of increasingly powerful machines with better-trained experts, this process allows businesses to produce more effectively, sustainably, and safely. Industry 5.0 is a paradigm shift in manufacturing that has implications for productivity, economics, and business. Therefore, companies that do not adapt their business model to the factory 5.0 model will quickly become obsolete, unable to benefit from the competitive advantages it offers. Towards that end, European Commission highlights the key features, the goals, the technological enablers and the challenges of Industry 5.0 in the framework presented in Fig. 16. Finally, the benefits of Industry 5.0 can be summarized as follows: a) Cost optimization, b) Personalization and Creativity and c) Greener and sustainable solutions [91].

7. CONCLUSIONS

BDA as a pillar technology of Industry 4.0 is becoming a core technology for forecasting and decision making in manufacturing systems towards the next Industrial Revolution (Industry 5.0). Big Data technologies consider not only the solution for large datasets, but also the ability to comprehend and fully exploit their value. BDA is a key future perspective in both the research and industrial communities, as it adds value to a variety of products and systems by incorporating cutting-edge technologies into traditional manufacturing products. Data can benefit from IoT, AI, and Industry 5.0 with built-in amendments to the design of future innovation ecosystems. Information technologies and manufacturing processes are brought together in Industry 4.0.

The production line may be tuned to near-optimum efficiency and even altered on the fly thanks to developments in robotics and automation, as well as in data collection and AI. Therefore, manufacturers are able to produce products and services of higher quality taking into consideration the following benefits [92]:

- Return of Investment (ROI),
- Higher revenue and profitability,
- Opportunities for innovation,
- Cost reductions,
- Improved customer experience,
- Compliance,
- Flexibility and Agility,
- Knowledge sharing and collaboration,
- Efficiency,
- Productivity.

Consequently, at this point it can be concluded from the above-mentioned points and the analysis of the corresponding research works, the fusion of the cutting-edge digital technologies (e.g. AI, robotics and automation, etc.) have enabled engineers to design resilient manufacturing and plants, which are capable of adapting and aligning their operating specifications/characteristics to the volatility of current market demands. By extension, these new type of manufacturing and production systems enables companies to handle more successfully market crises, such as the global pandemic caused by the SARS CoV-2 virus.

Similarly, Automation benefits can also be evident in the supply chain, as per the benefits below:

- In real-time, track the position of your goods,
- Track the status of your items in real time,
- Real-time data capture and analysis,
- Inventory/Warehouse Management: Ensure that the appropriate products are delivered to the right location at the right time.

In Industry 4.0, the abovementioned features are conceivable. However, the simplest way to define 5.0 is that it injects a conventional, individualized human touch into the automated and efficient concepts discussed before. Therefore, the benefits of Industry 5.0 in manufacturing can be summarized as follows [92]:

- Customization and personalization,
- Upskilling,
- Better roles for the human worker,
  - Using robots for repetitive and labor-intensive work,
  - Using humans for customization and thinking radically out of the box,
- Enhanced Customer Experience,
- Co-existence of humans and robots.

To summarize, based on the abovementioned challenges and case studies the quality of data is a major issue. However, with the combination of two major industry 4.0 technologies and techniques (AI, Ontologies) engineers are capable of fully connecting machine tools and creating things on the Internet (IIoT) in order to acquire meaningful insights for the manufacturing/production processes. On the other hand, in order to fully utilize this massive amount of data, engineers are developing frameworks for processing / cleaning raw data based on the integration of AI techniques. Moreover, ontologies can be realized as a technological field (specifically in Computer Science) with which engineers are capable of extracting context from raw data. By extension, the combination of the abovementioned technologies is believed to lay the foundations for acquiring the necessary amount of data (quantity of data) and simultaneously extract the informational context (quality of data). Ultimately, under the Industry 4.0 framework engineers had the opportunity to retrofit existing manufacturing/production systems with sensing systems, develop suitable network infrastructures (e.g. Wireless Sensor Networks – WSN) for fast and reliable transmission of data, as well as design and develop frameworks for monitoring and adjusting manufacturing process parameters. As a result, from our point of view Industry 4.0 is not in its infancy, there is however room for further improvement.

Finally, building complex and hyperconnected digital networks without compromising the long-term safety and sustainability of an innovation ecosystem and its constituents is the goal of Industry 5.0. In this paper, we summarized the state of the art in AI and semantic for Big Data analytics towards the 5th Industrial Revolution. Future research work as a continuation of this paper will be focused on the investigation of adding semantic in dynamic and interactive multi-objective algorithms for solving Big Data Optimization problems.

REFERENCES

[1] EUROSTAT, 2019, *Employment Specialisation of EU Regions,* Available at: https://ec.europa.eu/eurostat/web/-products-eurostat-news/-/DDN-20190917-1.

[2] CLARK D., 2021, *Number of Employees in Europe 2020,* Available at: https://www.statista.com/statistics-/1195197/employment-by-sector-in-europe/.

[3] LABOR MARKET BRIEFING SERIES, 2020, *The Manufacturing Sector in Europe,* Available at: https://cdn4.euraxess.org/sites/default/files/labor_market_information-_manufacturing_sector.pdf.

[4] LANZA G., FERDOWS K., KARA S., MOURTZIS D., SCHUH G., VÁNCZA J., WANG L., WIENDAHL H.P., 2019, *Global Production Networks: Design and Operation,* CIRP Annals, 68/2, 823–841, DOI: https://doi.org/10.1016/j.cirp.2019.05.008.

[5] CHRYSSOLOURIS G., 2006, *Manufacturing Systems: Theory and Practice (2nd ed.),* New York, NY, Springer Verlag.

[6] MOURTZIS D., 2011, *Internet Based Collaboration in the Manufacturing Supply Chain*, CIRP Journal of Manufacturing Science and Technology, 4/3, 296–304, DOI: https://doi.org/10.1016/j.cirpj.2011.06.005.

[7] LU Y., XU X., WANG L., 2020, *Smart Manufacturing Process and System Automation – A Critical Review of the Standards and Envisioned Scenarios,* Journal of Manufacturing Systems, 56, 312–325. DOI: https://doi.org/10.1016/j.jmsy.2020.06.010.

[8] MOURTZIS D., DOUKAS, M., 2014, *The Evolution of Manufacturing Systems: From craftsmanship to the Era of customization. Design and Management of Lean Production Systems,* IGI Global, 1–29, DOI: http://dx.doi.org/10.4018/978-1-4666-5039-8.

[9] YANG F., GU S., 2021, *Industry 4.0, a Revolution that Requires Technology and National Strategies,* Complex Intell. Systems, 7, 1311–1325, DOI: https://doi.org/10.1007/s40747-020-00267-9.

[10] MOURTZIS D., MILAS N., VLACHOU E., LIAROMATIS J., 2018, *Digital Transformation of Structural Steel Manufacturing Enabled by IoT-Based Monitoring and Knowledge Reuse,* International Conference on Control. Decision and Information Technologies CoDIT'18, 295–301, DOI: https://doi.org/10.1109/CoDIT.-2018.8394874.

[11] MOURTZIS D., MILAS N., ATHINAIOS N., 2016, *Towards Machine Shop 4.0: A General Machine Model for CNC machine-tools through OPC-UA*, Procedia CIRP 78, 301–306, DOI: HTTPS://DOI.ORG/10.1016/-J.PROCIR.2018.09.045.

[12] MOURTZIS D., VLACHOU E., 2018, *A Cloud-Based Cyber-Physical System for Adaptive Shop-Floor Scheduling and Condition-Based Maintenance*, Journal of manufacturing systems, 47, 179–198. DOI: https://doi.org/10.1016/j.jmsy.2018.05.008.

[13] MOURTZIS D., 2020, *Simulation in the Design and Operation of Manufacturing Systems: State of the Art and New Trends*, International Journal of Production Research, 58/7, 1927–1949, DOI: https://doi.org/10.1080/-00207543.2019.1636321.

[14] *Industry 5.0: Towards a Sustainable, Human-Centric and Resilient European industry*, DOI: https://doi.org/10.2777/308407.

[15] BENOTSMANE R., KOVÁCS G., & DUDÁS L., 2019, *Economic, Social Impacts and Operation of Smart Factories in Industry 4.0 Focusing on Simulation and Artificial Intelligence of Collaborating Robots,* Social Sciences, 8/5, 143, DOI: http://dx.doi.org/10.3390/socsci8050143.

[16] MOURTZIS D., 2020, *Machine Tool 4.0 in the Era of Digital Manufacturing*, Proceedings of the 32nd European Modeling & Simulation Symposium (EMSS 2020), 416–429, DOI: https://doi.org/10.46354/i3m.2020.emss.060.

[17] LIU C., VENGAYIL H., ZHONG Y.R., XU X., 2018, *A Systematic Development Method for Cyber-Physical Machine Tools*, Journal of Manufacturing Systems, 48, 13–24, DOI: https://doi.org/10.1016/j.jmsy.2018.02.001.

[18] ROMERO D., BERNUS P., NORAN O., STAHRE J., FAST-BERGLUND Å., 2016, *The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems,* IFIP

International Conference on Advances in Production Management Systems. Springer, Cham, DOI: https://doi.org/10.1007/978-3-319-51133-7_80.
[19] GAZZANEO L., PADOVANO A., UMBRELLO S., 2020, *Designing Smart Operator 4.0 for Human Values: A Value Sensitive Design Approach*, Procedia Manufacturing, 42, 219–226, DOI: https://doi.org/10.1016/-j.promfg.2020.02.073.
[20] GAO R., WANG L., HELU M., TETI R., 2020, *Big Data Analytics for Smart Factories of the Future*, CIRP Annals, 69/2, 668–92, DOI: https://doi.org/10.1016/j. cirp.2020.05.002.
[21] MOURTZIS D., DOUKAS M., PSAROMMATIS F., 2014, *Design of Manufacturing Networks for Mass Customisation Using an Intelligent Search Method*, International Journal of Computer Integrated Manufacturing, 28/7, 679–700, DOI: https://doi.org/10.1080/0951192X.2014.900867.
[22] MOURTZIS D., DOUKAS M., PSAROMMATIS F., 2015, *A Toolbox for the Design, Planning and Operation of Manufacturing Networks in a Mass Customisation Environment*, Journal of Manufacturing Systems, 36, 274–286, DOI: https://doi.org/10.1016/j.jmsy.2014.06.004.
[23] MOURTZIS D., DOUKAS M., PSAROMMATIS F., 2013, *Design and Operation of Manufacturing Networks for Mass Customisation*, CIRP Annals, 62/1, 467–470, DOI: https://doi.org/10.1016/j.cirp.2013.03.126.
[24] WANG G., GUNASEKARAN A., NGAI E.W.T., PAPADOPOULOS T., 2016, *Big Data Analytics in Logistics and Supply Chain Management: Certain Investigations for Research and Applications*, Int. J. Prod. Econ. 176, 98–110, DOI: https://doi.org/10.1016/j.ijpe.2016.03.014.
[25] KUSIAK A., 2017, *Smart Manufacturing Must Embrace Big Data*, Nature 544, 7648, DOI: https://doi.org/-10.1038/544023a.
[26] ZHONG R.Y., NEWMAN S.T., HUANG G.Q., LAN S., 2016, *Big Data for Supply Chain Management in the Service and Manufacturing Sectors: Challenges, Opportunities, and Future Perspectives*, Comput. Ind. Eng., 10, DOI: https://doi.org/10.1016/j.cie.2016.07.013.
[27] MOURTZIS D., GARGALLIS A., ZOGOPOULOS V., 2019, *Modelling of Customer-Oriented Applications in Product Lifecycle Using RAMI 4.0*, Procedia Manufacturing, 28, 31–36, DOI: https://doi.org/10.1016/j.promfg.2018.12.006.
[28] WEYRICH M., EBERT C., 2016, *Reference Architectures for the Internet of Things*, IEEE Software, 33/1, 112–116, DOI: 10.1109/MS.2016.20.
[29] FORREST C., 2016, *Big Data, Business Analytics to hit $203 Billion by 2020, Says IDC report, Available at:* https://www.techrepublic.com/article/big-data-business-analytics-to-hit-203-billion-by-2020-says-idc-report/.
[30] IBM, 2020, *Building a Robust, Governed Data Lake for AI*, Available at: https://www.ibm.com/-downloads/cas/RMAMZNRY.
[31] KURTZ J., SHOCKEY R., 2013, *Analytics: The Real-World Use of Big Data in Manufacturing*, IBM Global Business Services Business Analytics and Optimization, Available at: chrome-extension://-efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.ibm.com%2Fdownloads%2Fcas%2FONBGKB82.
[32] Big Data Market Size Revenue Forecast Worldwide from 2011 to 2027, Available at: https://www.statista.com/statistics/254266/global-big-data-market-forecast/.
[33] WANG L., 2019, *From Intelligence Science to Intelligent Manufacturing*, Engineering, 5/4, 615–8, DOI: https://doi.org/10.1016/j.eng.2019.04.011.
[34] RÜSMANN M., LORENZ M., GERBERT P., WALDNER M., et al., 2015, *Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries*, Boston, The Boston Consulting Group.
[35] MEERMANN L., 2019, *Sensors As Drivers of Industry 4.0: A study on Germany*, Switzerland and Austria, EY Building a better working world.
[36] DEMIR K. A., HALIL C., 2018, *The Next Industrial Revolution: Industry 5.0 and Discussions on Industry 4.0, Industry 4.0 from the Management Information Systems Perspectives*, Peter Lang Publishing House.
[37] SKOBELEV P., YU., B.S., 2017, *On the Way from Industry 4.0 to Industry 5.0: from Digital Manufacturing to Digital Society*, Industry 4.0, 2/6, 307–311.
[38] CISCO, 2020, *The Internet of Everything – Cisco IoE Value Index Study*, Available at: https://www.cisco.com/c/dam/en_us/about/business-insights/docs/ioe-value-index-faq.pdf.
[39] MOURTZIS D., VLACHOU E., MILAS N., DIMITRAKOPOULOS G., 2016, *Energy Consumption Estimation for Machining Processes Based on Real-Time Shop Floor Monitoring via Wireless Sensor Networks*, Procedia CIRP, 57, 637–642, DOI: https://doi.org/10.1016/j.procir.2016.11.110.

[40] MÜLLER J.P., FISCHER K., 2014, *Application Impact of Multi-agent Systems and Technologies: A Survey*, In: Shehory O., Sturm A., (eds) Agent-Oriented Software Engineering, Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-54432-3_3.

[41] MODONI G.E., DOUKAS M., TERKAJ W., SACCO M., MOURTZIS D., 2017, *Enhancing Factory Data Integration Through the Development of an Ontology: from the Reference Models Reuse to the Semantic Conversion of the Legacy Models*, International Journal of Computer Integrated Manufacturing, 30, 1043–1059. DOI: https://doi.org/10.1080/0951192X.2016.1268720.

[42] MALEKI E., BELKADI F., BOLIET N., al., 2018, *Ontology-Based Framework Enabling Smart Product-Service Systems: Application of Sensing Systems for Machine Health Monitoring*, IEEE Internet of Things Journal, 5/6, 4496–4505, DOI: https://doi.org/10.1109/JIOT.2018.2831279.

[43] RZEVSKI G., SKOBELEV P., 2007, *Emergent Intelligence in Large Scale Multi-Agent Systems*, International Journal of Education and Information Technology, 1/2, 64–71.

[44] VITTIKH V., 2016, *Evergetics: Science of Intersubjective Management Processes in Everyday Life*, International Journal of Management Concepts and Philosophy, 9, 63, DOI: https://doi.org/10.1504/IJMCP.2016.077763.

[45] TAOULI A., DJAMEL A.B., KESKES N., BENCHERIF K., HASSAN B., 2018, *Semantic for Big Data Analysis: A survey*, NTIS2018, BigData & Internet of things IoT.

[46] FRICKÉ M.H., 2018, *Data-Information-Knowledge-Wisdom (DIKW) Pyramid, Framework, Continuum*. In: Schintler L., McNeely C., (eds) Encyclopedia of Big Data, Springer, Cham, DOI: https://doi.org/10.1007/978-3-319-32001-4_331-1.

[47] YAO X., LIAN Z., YANG Y., ZHANG Y., JIN H., 2014, *Wisdom Manufacturing: New Humans-Computers-Things Collaborative Manufacturing Model*, Computer Integrated Manufacturing Systems, 20/6, 1490–1498. DOI: http://dx.doi.org/10.13196/j.cims.2014.06.yaoxifan.1490.9.20140627.

[48] OXFORD LEARNER'S DICTIONARY, Available at: https://www.oxfordlearnersdictionaries.-com/us/definition/english/wisdom?q=WISDOM.

[49] ROWLEY J., 2007, *The Wisdom Hierarchy: Representations of the DIKW Hierarchy*, Journal of Information Science, 33/2, 163–180, DOI: https://doi.org/10.1177/0165551506070706.

[50] ZHENG X., HU Y., XIE K., ZHANG W., SU L., LIU M., 2015, *An Evolutionary Trend Reversion Model for Stock Trading Rule Discovery*, Knowledge-Based Systems 79, 27–35. DOI: https://doi.org/10.1016/j.knosys.-2014.08.010.

[51] KHODKH P., LAWANGE S., BHAGAT A., DONGRE K., INGOLE C., 2010, *Query Processing Over Large RDF Using Sparql in Big Data*, Communications of the ACM, DOI: http://dx.doi.org/10.1145/-2905055.2905124.

[52] LEIDA M., RUIZ C., CERAVOLO P., 2016, *Facing Big Data Variety in a Model Driven Approach*, IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), 1–6, DOI: https://doi.org/10.1109/RTSI.2016.7740641.

[53] KELLER R.M., RANJAN S., WEI M.Y., ESHOW M.M., 2016, *Semantic Representation and Scale-up of Integrated Air Traffic Management Data*, Proceedings of the International Workshop on Semantic Big Data, 1–6. DOI: https://doi.org/10.1145/2928294.2928296.

[54] ISSA H.,L. ELST V., DENGEL A., 2016, *Using Smartphones for Prototyping Semantic Sensor Analysis Systems*, SBD'16, Proceedings of the International Workshop on Semantic Big Data 7, 1–6, DOI: https://doi.org/10.1145/2928294.2928299.

[55] KAISLER S., ARMOUR F., ESPINOSA J., MONEY W., 2013, *Big Data: Issues and Challenges Moving Forward*, 46th Hawaii International Conference on System Sciences (HICSS), 995–1004, DOI: http://dx.doi.org/10.1109/HICSS.2013.645.

[56] ZHANG X., ZHAO J., LECUN Y., 2015, *Character-Level Convolutional Networks for Text Classification*, Proceedings of the 28th International Conference on Neural Information Processing Systems, 1, 649–657.

[57] NURAL M., COTTERELL M., MILLER J., 2015, *Using Semantics in Predictive Big Data Analytics*, IEEE International Congress on Big Data, 254–261, DOI: https://doi.org/10.1109/BigDataCongress.2015.43.

[58] BANAGE T.G., S. KUMARA I., PAIK J., ZHANG T.H.A., SIRIWEERA S., KOSWATTE K.R.C., 2015, *Ontology-Based Workflow Generation for Intelligent Big Data Analytic*, IEEE International Conference on Web Services, 495–502, DOI: https://doi.org/10.1109/ICWS.2015.72.

[59] YAO Y., ZHANG L., YI J., PENG Y., HU W., SHI L., 2016, *A Framework for Big Data Security Analysis and the Semantic Technology*, 6th International Conference on IT Convergence and Security (ICITCS), 1–4. DOI: https://doi.org/10.1109/ICITCS.2016.7740303.

[60] NURAL M.V., COTTERELL M.E., PENG H., XIE R., MA P., MILLER J., 2015, *Automated Predictive Big Data Analytics Using Ontology Based Semantics*, International Journal of Big Data 2/2, 43–56, DOI: https://doi.org/10.29268/stbd.2015.2.2.4.

[61] NAJAFABADI M.M., VILLANUSTRE F., KHOSHGOFTAAR T.M., SELIYA N., WALD R., MUHAREMAGIC E., 2015, *Deep Learning Applications and Challenges in Big Data Analytics,* Journal of Big Data, 2/1, 1–21, DOI: https://doi.org/10.1186/s40537-014-0007-7.

[62] BENGIO Y., COURVILLE A., VINCENT P., 2013, *Representation Learning: A Review and New Perspectives*, Pattern Analysis and Machine Intelligence, IEEE Transactions, 35/8, 1798–1828. DOI: https://doi.org/10.1109/TPAMI.2013.50.

[63] BENGIO Y., LECUN Y., 2007, *Scaling Learning Algorithms Towards AI*, In: Bottou L, Chapelle O., DeCoste D., Weston J., (eds) Large Scale Kernel Machines, MIT Press, Cambridge, MA, 34, 321–360, Available at: http://www.iro.umontreal.ca/~lisa/pointeurs/bengio+lecun_chapter2007.pdf.

[64] HINTON G., SALAKHUTDINOV R., 2010, *Discovering Binary Codes for Documents by Learning Deep Generative Models*, Topics Cogn. Sci. 3/1, 74–91, DOI: https://doi.org/10.1111/j.1756-8765.2010.01109.x.

[65] RANZATO M, SZUMMER M., 2008, *Semi-Supervised Learning of Compact Document Representations with Deep Networks*, Proceedings of the 25th International Conference on Machine Learning, ACM, 792–799.

[66] MIKOLOV T., CHEN K., DEAN J., 2013, *Efficient Estimation of Word Representations in Vector Space*, CoRR: Computing Research Repository, 1–12.

[67] CHENG S., SHI Y., QIN Q., BAI R., 2013, *Swarm Intelligence in Big Data Analytics*, International Conference on Intelligent Data Engineering and Automated Learning, Springer, 417–426.

[68] ABDUL-RAHMAN S., BAKAR A.A., MOHAMED-HUSSEIN Z.A., 2013, *Optimizing Big Data in Bioinformatics with Swarm Algorithms*, Computational Science and Engineering (CSE), 1091–1095, DOI: https://doi.org/10.1109/CSE.2013.158.

[69] GOVINDARAJAN K., SOMASUNDARAM T.S., KUMAR V.S., et al., 2013, *Continuous Clustering in Big Data Learning Analytics*, Technology for Education (T4E), 2013 IEEE Fifth International Conference, 61–64, DOI: https://doi.org/10.1109/T4E.2013.23.

[70] TANNAHILL B.K., JAMSHIDI M., 2014, *System of Systems and Big Data analytics, Bridging the gap,* Computers and Electrical Engineering, 40/1, 40th-year commemorative issue, 2 –15.

[71] CABANAS-ABASCAL A., GARCÍA-MACHICADO E., PRIETO-GONZÁLEZ L., DE AMESCUA SECO A., 2013, *An Item based Geo-Recommender System Inspired by Artificial Immune Algorithms*, J. Univers. Comput. Sci., 19, 2013–2033.

[72] LEE W.-P., HSIAO Y.-T., HWANG W.-C.., 2014, *Designing a Parallel Evolutionary Algorithm for Inferring Gene Networks on the Cloud Computing Environment,* BMC Systems Biology, 8/1, 1–19.

[73] SUN W., ZHANG N., WANG H., YIN W., QIU T., 2013, *PACO: A Period ACO Based Scheduling Algorithm in Cloud Computing*, Cloud Computing and Big Data (CloudCom-Asia), 2013 International Conference on 482–486.

[74] BARBA GONZÁLEZ C., 2018, *Big Data Optimization: Algorithmic Framework for Data Analysis Guided by Semantics*, PhD thesis, University of Malaga.

[75] NEBRO A.J., RUIZ A.B., BARBA-GONZÁLEZ C., GARCÍA-NIETO J., LUQUE M., ALDANA- MONTES J.F., 2018, *InDM2: Interactive Dynamic Multi-Objective Decision Making Using Evolutionary Algorithms,* Swarm and Evolutionary Computation 40, 184–195.

[76] NOY N.F., MCGUINNESS D.L., 2001, *Ontology Development 101*: *A Guide to Creating your First Ontology*, Available at: https://protege.stanford.edu/publications/ontology_development/ontology101.pdf.

[77] STAAB S., STUDER R., 2003, *Handbook on Ontologies*, Springer Science & Business Media, DOI: https://doi.org/10.1007/978-3-540-92673-3.

[78] MCGUINNESS D.L. VAN HARMELEN F., 2004, *OWL Web Ontology Language Overview*, W3C Recommendation 10.10.

[79] HARRIS S., SEABORNE A., PRUD'HOMMEAUX E., 2013, *SPARQL 1.1 Query Language*, W3C Recommendation 21.10.

[80] GROSOF B.N. POON T.C., 2004, *SweetDeal: Representing Agent Contracts with Exceptions Using SemanticWeb Rules, Ontologies, and Process Descriptions,* Int. Journal of Electronic Commerce, 8/4, 61–97.

[81] COMPTON M., BARNAGHI P., BERMUDEZ L., GARCIA-CASTRO R., et al., 2012, *The SSN Ontology of the W3C Semantic Sensor Network Incubator Group,* Web Semantics: Science, Services and Agents on the World Wide Web, 17.

[82] MOURTZIS D., VLACHOU E., MILAS N.J.P.C., 2016, *Industrial Big Data As a Result of IoT Adoption in Manufacturing*, Procedia CIRP, 55, 290–295, DOI: https://doi.org/10.1016/j.procir.2016.07.038.
[83] ZHANG J., WANG J., LYU Y., BAO J., 2019, *Big Data Driven Intelligent Manufacturing*, Zhongguo Jixie Gongcheng/China Mech. Eng., 30/2, DOI: https://doi.org/10.3969/j.issn.1004-132X.2019.02.001.
[84] GLIMM B., STUCKENSCHMIDT H., 2016, *15 years of semantic web: An incomplete survey*, KI-Künstliche Intelligenz, 30/2, 117–130.
[85] TAYLOR A., 2015, *Semantics for Dummies*, John Wiley & Sons, Inc.
[86] BAUMGAERTEL H., 2017, *Semantic Web in Industrial Companies - Methods, Architectures, Applications*, DOI: 10.13140/RG.2.2.15745.56166.
[87] WANG J., XU C., ZHANG J., ZHONG R., 2021, *Big Data Analytics for Intelligent Manufacturing Systems: A review*, Journal of Manufacturing Systems, DOI: https://doi.org/10.1016/j.jmsy.2021.03.005.
[88] TILBURY D.M., 2019, *Cyber-Physical Manufacturing Systems*, Annual Review of Control, Robotics, and Autonomous Systems, 2, 427–443, DOI: https://doi.org/10.1146/annurev-control-053018-023652.
[89] CULOT G., FATTORI F., PODRECCA M., SARTOR M., 2019, *Addressing Industry 4.0 Cybersecurity Challenges*, IEEE Engineering Management Review, 47/3, 79–86, DOI: 10.1109/EMR.2019.2927559.
[90] LEZZI M., LAZOI M., CORALLO A., 2018, *Cybersecurity for Industry 4.0 in the current literature: A reference framework*, Computers in Industry, 103, 97–110, DOI: https://doi.org/10.1016/j.compind.2018.09.004.
[91] MÜLLER J., DIRECTORATE-GENERAL for RESEARCH and INNOVATION (EUROPEAN COMMISSION), 2020, *Enabling Technologies for Industry 5.0*, DOI: https://doi.org/10.2777/082634.
[92] DALENOGARE L.S., BENITEZ G.B., AYALA N.F., FRANK A.G., 2018, *The Expected Contribution of Industry 4.0 Technologies for Industrial Performance*, International Journal of Production Economics, 204, 383–394. DOI: https://doi.org/10.1016/j.ijpe.2018.08.019.