Industry Focused in Data Collection

How Industry 4.0 is Handled by Big Data

Miguel Oliveira
School of Design, Management and Production
Technologies Northern Aveiro,
University of Aveiro
Oliveira de Azeméis, Aveiro, Portugal
miguel@ua.pt

Daniel Afonso
School of Design, Management and Production
Technologies Northern Aveiro,
University of Aveiro
Oliveira de Azeméis, Aveiro, Portugal
dan@ua.pt

ABSTRACT

The paper aims to organize and structure data collected and associated to technologies that powers the abroad concept of Industry 4.0. It starts with the historic evolution of industry, separated by date landmarks and approaches the last transition between 3.0 to 4.0.

Apart from the differences between industry models, production data stats show a huge and important transformation in the amount of data related to manufacturing and how that knowledge is processed. The paper also aims to put on debate the lack of solutions regarding the knowledge extraction of data from machines and systems, needed for data analytics.

Approaches with cyber-physical systems, machine learning, virtual environments, Industrial IoT ¹ and augmented reality, in an industrial scale, are some of the strategies to power the reading and interpretation of data, in order to promote industrial efficiency.

Real context industrial applications are taken in account in order to state the importance of collected data in the efficiency of a production process. Exploring technologies and concepts to improve digital twins systems, perception and perceived systems as well as maintenance processes are some of the explored implemented strategies that make Industry 4.0. Some possible strategies are presented, as well as the transition for Industry 5.0.

CCS CONCEPTS

• Applied computing • Operations research • Industry and manufacturing

KEYWORDS

Industry 4.0, Data analytics, Big data, Knowledge data, IoT, CPS, Smart Factory

ACM Reference format:

1 A Brief of Industrial History

Being the production of goods or services within an economy[1] and the economic activity concerned with the processing of raw materials and manufacture of goods in factories[2] or, yet, the group of companies that are related based on their primary business activities[3], the industrial model has had four major changes in the last four centuries. The first revolution occurred between 1760 to 1840 with the development of machines[4], mechanical production equipment, mainly used for textile industry powered by water and steam power[5]. The second revolution dates from the last quarter of the 19th century to the first quarter of the 20th century[6], also known as the technological revolution, characterized by the introduction of electrical energy, motors and division of labor in mass production systems. The third revolution emerged in the 70’s of 20th century, also known as the digital revolution[7], where the electronics/IT to automate production has had an important role. The Industry 4.0 was named for the first time in 2011[8], also referred as “smart factory”, and relies on cyber-physical systems, IoT, big data and artificial vision[9], where interoperability and intercommunication between technologies and systems is made to help professionals and systems to make decisions. Figure 1 illustrates those transitions over the centuries.
1.1 Industry 4.0

Under the initiative of the German federal government in its High-Tech Strategy for 2020, the Industry 4.0 term was first time used in 2011 [11] in Hannover Fair, becoming a landmark for the separation with the previous concept (3.0). The outcome they pretend to foster states that:

"In the future, businesses will establish global networks that incorporate their machinery, warehousing systems and production facilities in the shape of Cyber-Physical Systems (CPS). In the manufacturing environment, these Cyber-Physical Systems comprise smart machines, storage systems and production facilities capable of autonomously exchanging information, triggering actions and controlling each other independently. This facilitates fundamental improvements to the industrial processes involved in manufacturing, engineering, material usage and supply chain and life cycle management. The Smart Factories that are already beginning to appear employ a completely new approach to production. Smart products are uniquely identifiable, may be located at all times and know their own history, current status and alternative routes to achieving their target state. The embedded manufacturing systems are vertically networked with business processes within factories and enterprises and horizontally connected to dispersed value networks that can be managed in real time – from the moment an order is placed right through to outbound logistics. In addition, they both enable and require end-to-end engineering across the entire value chain."[11]

The Industry 4.0 is based on several technologies that help factories to take better decisions in order to boost productivity and efficiency at several layers. It is supported by collecting and transmitting data from and between machines, using sensors, artificial vision, embedded electronics and Industrial IoT, among others. These major technology advances comprehend the cyberphysical production systems, producing huge amounts of data that must be processed, converted in knowledge and, therefore, learned.

The major difference from the 3.0 to 4.0 evolution, comparing to the other two transitions (1.0 to 2.0 and 2.0 to 3.0) regards the fact that this is more affordable. It’s widely accepted that robots, cameras, sensors and other devices are available at a distance of click from retailers at low range cost. From recent years, IoT devices prices have been dropping, becoming cheaper and more featured, allowing small or big companies to purchase them. This opens a wide range of opportunities to explore in terms of collecting data. Stats picture this scenario.

1.2 Cloud and analytics

In a recent GE2 whitepaper [12], it is emphasized how industrial data has becoming industrial big data, where the amount of data generated is getting bigger each day. As an example, the report states that in a single machine in a baby product manufacturing process, it is generated 152K data samples per second, 9M data per minute, 545M per hour, 4B samples per shift, 13B samples per day and 4T samples per year [12]. This stats regards only one machine, where more data like geolocation, errors, warnings and operation logs may be added. Thus, the increase amount of data related to machines is something that is exponentially growing. The question that remains is how that data is handled and how to carry out knowledge from data in a fast and smart way?

Coming from different sources and in different formats, there is a huge need to process this data, convert it in knowledge and use it in a proactive and predictive manner, creating knowledge value and building digital trust. Recent research shows that data-driven decision-makers demonstrate 4% higher productivity, 6% greater profitability, and 50% higher market value from Information Technologies (IT) [13]. Although the path is being made in attempt to interoperate data from different machines, this lead us to the ambition of objectiveness on how to put all together in a proper mean to interpret the data. In a 2015 working paper, Hermann, Pentek, and Otto set the design principles of Industry 4.0, identifying them in the Industry 4.0 main components [14].

1.3 Industry 4.0 Components

As been said in prior section, Hermann, Pentek, and Otto identify two main components detailed by the “Industrie 4.0 Working Group”: Cyber-Physical Systems (CPS) and Smart Factory.

CPS comprehend computer algorithms that manages and monitors mechanisms or machines that interoperate with users and Internet. In CPS, physical and software components are deeply intertwined, each operating on different spatial and temporal scales, exhibiting multiple and distinct behavioral modalities, and interacting with each other in a lot of ways that change with context [15]. Some examples are the car factory automation processes, aviation cruise controls, robots, etc. The term CPS is prior to the new transition of Industry but represents and important component. The major reason is because it allows the creation and implementation of a digital twin with the data available, as well as the ability of control it. A digital twin, also known as Background Systems [17], is a congregated model of a physical and virtual machine, which enables to set several actions and production scenarios in order to prevent future problems and to increase productivity and efficiency. This combined concept (real and virtual) gives an important role of the CPS component.

Although broad, the definition of Smart Factory is a concept that emerged in recent years due to the assortment between big data,

---

2 General Electric
that allows the analysis of complex processes and information to refine production, advanced robotics controlled by machines and algorithm and, also, interconnectivity and interoperability between devices and services in order to acknowledge best production performances. A Smart Factory is defined as a factory that context-aware assists people and machines in the execution of their tasks [14] and constitute a key feature of Industry 4.0[16]. This component promotes best practices in factory workplace due to the integration of every production and development process [18]. Also, tends to reduce inefficiency and vulnerabilities, powering safety [19], recurring, as an example, to the development of simulations to obtain major, safety and better results, from problem to solution.

The other two components referred by Hermann, Pentek, and Otto are the IoT and Internet of Services (IoS). Although these components are not explicitly mentioned in the report of Industrie 4.0 Working Group, they are implicitly related with the previous components as they act as hub or network stakeholder to materialize the Industry 4.0 concept.

IoT empowers the use of Internet technologies in things and objects such as RFID, sensors, actuators, mobile phones, which, through unique addressing schemas, (...) interact with each other and cooperate with their neighboring ‘smart’ components, to reach common goals [20]. This core technology gives the ability to monitor, control and manage different set of devices, enabling centralization of services and, thus, data in a real-time analysis, allowing the development of new business models that contribute to efficient production processes. According to Patel [12], in the last 10 years, the cost of sensors dropped 0.60 times, RAM, 21 times, processing, 60 times, disk storage, 17 times, bandwidth, 40 times and flash memories 646 times. This is a huge drop scale that represents the suppression of a problem regarding technology infrastructures and endorsed the development of Industrial IoT, i.e., the enforcement of IoT on Industrial models.

IoS, Cloud Services, or Cloud Computing, stands for services offered through internet channels. This definition aids users to release devices from centralized software, powering connections between them, and simplifying the process of software maintenance, installation and integration. It also promotes the centralization of data in one place, rather than spread it in several equipment and devices. It is conceivable that this concept will be transferred from single factories to entire value-added networks in the future. Factories may go one step further and offer special production technologies instead of just production types. These production technologies will be offered over the IoS and can be used to manufacture products or compensate production capacities [21].

The recognition of these components is based on the research made by the authors, regarding the analysis made through occurrences of these components terms in industry 4.0 related publications within a statistical universe of 51 publications [14].

1.4 Design Principles

The assumption of the previous components has driven manufacturers do identify 6 principles to be at stake when planning the industry transition strategy. These principles, according with authors [14], have been derived from the identified components in section 1.3. They are Interoperability, Virtualization, Decentralization, Real-Time Capability, Service Orientation and Modularity.

As the name says, interoperability is a concept that enables de communication between CPS and Smart Factory through IoT and IoS by standardized communication layers, protocols, semantics and data exchange formats. In our understanding, interoperability is the most important principle to promote the current industry transition as the core to connect every component.

Virtualization means that every CPS or CPS networks must be able to virtualize, simulate and emulate, creating a virtual copy of the real world. If a factory is able to virtualize every system and device in the workshop, generating data and real-context scenarios, through machine learning and complex algorithms, it will be able to produce knowledge to get the most out from machines, powering efficiency, thus, productivity, predicting failures, errors and warnings.

The increase of machines and devices in the workshop makes, in a disadvantage way, the centralization of everything harder. In order to suppress this problem, decentralization is a must for this transition. It means that every device, machine, CPS must be able to take action, i.e., decentralized decisions for problem solving in a predictive and preventive way. As a consequence, for making decisions on their own, these systems must be as flexible as they can, must ensure quality of service. Only in cases of failure tasks are delegated to a higher level [22].

Data collected (and stored), processed and analyzed in real-time is a demand for this new industry paradigm. Smart factories must be able to collect all the data they can, both to make the best decisions and to take the best actions and to analyze eventual faults and warnings to prevent production failure. Again, machines must be able to re-delegate tasks to other machines, or to request instruction set by a higher level or layer. Real-Time capability, apart of principle definition, stands for high quality technology solutions.

Service-orientation in production strategy implies the delivery of a product or service, having the customer in the center of demand and decision. This principle is essential for the cooperation between different industries in joint added value creation. The IoS gets an important role in this principle, since people, nowadays, are internet-connected, playing an important part in market creation, as well as product specification, as independent individuals or masses.

The last, although not less essential principle, stands for modularity. It comprehends that every part of a CPS, system and device are flexible for implementing different requirements, replacements or expansion modules. Modular systems in smart factories must be easily re-adapted when the demand of a product changes due to different target requirements, features or characteristics. Nowadays, derived by the fast changing market, decision makers and stakeholders must be capable to change
production smoothly and within a short range time window. Smart sustainable manufacturing processes able to operate with low volume production while economically viable are also a key feature for a smart factory.

Maybe add a short paragraph about green principles associated with new production technologies.

1.5 Gathering Principles and Components

In the previous section it has been said that interoperability is the most important key for this industry transition. Indeed, this is true because it is a mandatory requirement for all components (CPS, IoT, IoS) in which the smart factories are created. Allowing the connection and data exchange between components, making low and middle level decisions produce better data for a smart factory high level decision. If factories are able to create their digital twins, in each specific CPS and, as a whole, within a smart factory, they are promoting a predictive maintenance and efficient management of production, reducing costs and, most probably, variables on the production processes. In the same way, decentralized and modular systems allow to distribute hierarchized decisions from CPS, individually, and, globally, by smart factories. This endorses the ability to take real-time decisions and fast designing strategies for a smart factory, where all the amount of data may be processed and turned into knowledge in short time ranges for decision makers, in terms of production. Centered on customers and market trends, IoS stimulate the decision makers to plan, customize or adapt production oriented for individual aims, regarding specificities, features and desired products, as well as mass customization. This is an important step where the personalized production (customer specific production) has replaced craft production.

2 Industry Results

According to what is being said in the previous sections, the results are seen in different locations and industries. Probably, in one of the most advanced industries, the automobile, intelligent manufacturing environments are a reality, as well as the state of the art, where CPS coexist in large scale and fully network connected in factories. The factory virtualization, supported by digital twins networks, must be highly efficient, in desired outcomes, energy management and shop floor management. As an industrial example, the BMW Group (Fig.2) uses its Smart Data Analytics digitalization cluster to analyze this data selectively and enhance its production system [23]. Big data analysis contributes for an effectively quality improvement on production results, as well as logistics.

Figure 2: Robots and Control Technology are Fitted with the Necessary Sensors, Allowing Maintenance Staff Analyze the Data and Draw the Right Conclusions [24].

Data-driven improvements to processes and systems help reducing lead time and lower costs [23]. Using an intranet-of-things platform, that congregates huge amount of data from sensors and data from devices and systems, delivery real-time, the speed of the solution developed is considerable, having reduced the complexity and implementation costs in production and logistics, as well as the improvement of the shop floor management.

Other important sector that is committed to incorporate the new industry transition principles and components is the energy sector, in particular case the renewable energy sector. Using IoT sensors, wind energy companies analyze and adapt windfarms (Fig.3) according potential sites with high gross power in order to plan de location and configuration of wind farms [25].

Figure 3: Wind Farm[30]

As a result of a predictive maintenance, rather than preventive, thermal cameras are used to analyze wind turbines. This approach leads to huge data generation, stored and available for a thermography database. With complex the data, algorithms and machine learning for fault detection, companies may predict the eventual failure of a turbine in order to compensate de energy production with other systems. The research literature regarding this specific scenario is considerable, with several groups around the globe dedicated to the creation of new mathematical models, complex algorithms and machine learning for fault detection.
Having become one of the most important logistics facilitators to support the export of services, products and goods, the port of Rotterdam (Fig.4) is one of the most dynamic in the world in its area. Innovation is a must do, providing intelligent technologies in navigation assistance, ship tracking and containers parking and storage.

One of the most interesting service and technology provided by the company is a shared platform for information exchanges that relate to port calls [26].

Figure 4: Port of Rotterdam[31]

Involving all the players in the shipping activity (companies, dispatch agencies, logistic and maritime providers, authorities), through platform, app and api for data exchange, they enable the optimization of planning, execution and monitoring specific port activities, empowering efficiency, time reduction and smarter decisions.

One of the major mold production groups in the world, the Simoldes Group³, in Portugal, through a company called Inplás⁴, supported by the European Regional Development Fund⁵, have launched the “4.0 Excellence in Industrial Competitiveness and Innovation”. Throughout this project, Inplás aims to improve its installed capacity for process and organizational innovation, as well as the expansion of its product offer, associated with a set of automatisms and connections between equipment, software and people that allow the increase of efficiency, tending to reduce redundant tasks, excessive use of machine downtime, anomalies and rejections, and aimed at increasing the final quality, safety, accuracy, customization and flexibility, converging to a better time-to-market and an innovative / differentiated offer.

Analyzing the given examples it is clear that the principles, components, as well as technology options, where software is at the core of this revolution, aim to optimize and improve common, although specific, issues that are transversal to industries: smartness, fastness, productivity, energy and power use.

3 Core-manufacturing/Automation Players

This section addresses some of the main core-manufacturing/automation players that are stepping ahead of this needs and act as technology partners. Striving towards Industry 4.0, big companies are presenting new solutions, artificial entities that integrate autonomously between them and also with humans, to perform planned functions and tasks, intelligently, focusing on big data dependency to take decisions using cyber-physical systems.

GE⁶, as a major player on this market competition, disclosures a set of integrated solutions like application software for predictive manufacturing, plant monitoring, and intelligent production management. Other segments explore predictive analysis, complemented with machine learning and cyber security solutions for industrial IoT implementation. The solutions presented available tend to promote interoperability between applications and cloud, generating big amount of data.

Another important key player, Siemens⁷ emphasizes the manufacturing processes, as the core of their main solutions. They provide software and hardware solutions that allow a full scale integration in several industries, to manage big data in the complete production lifecycle. This specific area of the company is segmented in technologies like automation, control, software and consulting services.

Rockwell Automation⁸ have its main focus on intelligent production control, operation and monitoring solutions. They offer a wide range of intelligent devices connected to help in fast and best decision to make at the shop floor. Philosophy is based in operational intelligence, promoting efficient decisions in acceptable time windows, productivity enhanced by interoperability of control systems and also control risk management, recurring to predictive analysis.

Targeted specifically for IoT, Bosch⁹ centers its main portfolio on IoT and associated solutions for business and process management (BPM and BRM), and interoperability solutions for heterogeneous business. IoT plays the most important part in data generation although the integration of devices on business (not only in factories) act as gateway middleware for intelligent decision making.

As an historical industrial partner, Festo¹⁰ is dedicated to IoT for original equipment manufacturers (OEM). They stand for integrated solutions between the production and digital scenarios, in order to promote flexibility, energy saving and linking production to logistics, with a set of software and hardware.

There are also some companies like Cisco¹¹ or SAP¹² that fuel their main products, hardware and software, with services for integration with other partners or technologies. The most common conclusions from the potential partners are the investment in areas like augmented reality to support knowledge creation and dissemination, virtual reality for

---

10. https://www.cisco.com
simulation and emulation for predictive maintenance and scenarios, software solutions from single devices to cloud computing, with different hierarchy layers. Also, in the searches made, IoT cyber security is a huge concern as they offer solutions to prevent production, logistics and business threads. Although not presented with the conceptual name of “digital twins”, solutions for CPS are available in most of the examples referred previously.

4 Final Considerations

The current industry scenario, brought by the introduction of information technology into production, services, and all sectors of the economy, raises big challenges for companies, households and the public sector. The impact of change is so critical that it was referred, as been said before, as the Fourth Industrial Revolution, or Industry 4.0, where the core of this transition the industry production, supported by predictive maintenance, life cycle optimizations, sustainability promotion, as well as the digitalization and production networking, aiming to increase the economic growth. This evolution in the industry organization go along with new production concepts based on interoperability such as flexible manufacturing system (FMS), manufacturing on demand (MOD) and mass customization. As a consequence, professional activities must be endorsed to explore capable workers to deal with software development, robotics, control systems and automation. The US Bureau of Labor Statistics has disclosure a list with Fastest Growing Occupations for the 2016–2026 period, ranked by growth rate and year wage salary [27]. The two biggest growth stands for solar photovoltaic installers and wind turbine service technicians in which is, by the way, two major industries at the top of the head of the current industry transition. Repetitive, cycle and recurring tasks will no long be made by humans but by machines, with an increasingly efficient use, able to optimize the tasks and decisions they take. It is also understood in the ranked list that most of the occupations are related with sustainability issues, like bicycle repairers and forest fire inspectors. Other occupations are software developers and information security analysis. The referred occupations are engaged, directly or indirectly with industry issues and the need to update occupations skills and abilities.

A natural selection process has been happening while this transition occurs. The willingness to change, upgrade and adapt assets, in special case for human resources is a demand, with continuous professional development. Also, the mistake value recognition is a demand towards the development of knowledge databases that assists, as an example, digital twins with machine learning aiding. Companies that don’t assist and issue these strategies tend to lose the track with their efficient industry processes that turn into undesired results. Unfortunately, they also often fail to reflect the organization’s actual processes or to meet the real needs of manufacturing companies. Consequently, there are limited examples of transformational changes either within individual enterprises or at the level of the economy as a whole [28].

5 What’s Next? Industry 5.0, Present Time.

Phill Cartwright, executive chairman of the Centre for Modelling & Simulation, states that Industry 4.0 starts to move towards Industry 5.0 when you begin to allow customers to customize what they want [29]. It will allow customers to select personalized products from an endless set of options (or variables), enabling users to add value and finding new paths for product development. This can be seen as a conceptual model for the next transition, comprehending most of production life and business cycles, where the ‘maturation’ of industry 4.0 is leading towards to economic, social and ecological gains. The term Industry 5.0 has been firstly introduced on December 1st 2015 in article published in linkedin social network by Michael Rada. As a technical concept, this transition enables the connection between machines, i.e., machines communicating and sharing data with other machines and also with humans, where production is constantly optimized by machine learning and computing algorithms, through considerable amounts of data. But, rather than separated concepts, it is our understanding that the current 4.0 and 5.0 transitions are not so distinct from each other. The 4.0 transition have become with the computing of huge amounts of data, suppressing human decisions, helped mainly with CPS, that lead to short amount of time to the improvement of technical equipment and software where, now, is being referred on 5.0 transition.

REFERENCES


14 https://www.linkedin.com/pulse/what-industry-50-de-marcell-vollmer/
Industry Focused in Data Collection: How Industry 4.0 is Handled by Big Data

[24] Photo courtesy BMW Group
[28] Schuh, Günther, Anderl, Reiner, Gausemeier, Jürgen , Hompel, Michael ten, Wahlster, Wolfgang (2017), Industrie 4.0 Maturity Index, Managing the Digital Transformation of Companies, acatech STUDY.
[29] Rossi, Ben (March 7th, 2018), Manufacturing gets personal in Industry 5.0, Raconteur, https://www.raconteur.net/technology/manufacturing-gets-personal-industry-5-0