Process Mining towards Industry 4.0 maturity
Abstract

We are currently living the 4th Industrial Revolution and it is becoming complex to set a finish line since it embraces a really broad range of topics and features. In this context, companies are aware about the benefits of the Industry 4.0 and try to achieve them. Whereas some organizations attempt to define a roadmap towards maturity in Industry 4.0, it is not an easy task to assess a progression in real businesses.

New technologies and methodologies are needed to undertake contemporary challenges, and Process Mining emerges among others as a technique that finds its perfect spot in the Industry 4.0 scenario. Some of the top global companies have already started using it but there is still plenty of development left to implement.

Keywords

Industry 4.0 | Process Mining | Digitalization | Digital Twin | Manufacturing
# Table of contents

**Introduction** .......................................................................................................................... 6  
Motivation .................................................................................................................................. 7  
Problem statement ......................................................................................................................... 7  
Objectives ..................................................................................................................................... 8  

1. **Industry 4.0** ............................................................................................................................. 9  
   1.1. Definition ................................................................................................................................. 9  
      1.1.1. Benefits ............................................................................................................................... 11  
      1.1.2. Challenges ......................................................................................................................... 13  
      1.1.2.1. Industry 4.0 maturity ...................................................................................................... 14  
   1.2. Industrial Revolutions in history .............................................................................................. 15  
      1.2.1. First Industrial Revolution ............................................................................................... 16  
      1.2.2. Second Industrial Revolution ......................................................................................... 16  
      1.2.3. Third Industrial Revolution .............................................................................................. 17  
      1.2.4. Fourth Industrial Revolution ............................................................................................ 17  
   1.3. Industry 4.0 around the world ............................................................................................... 17  
      1.3.1. German industry .............................................................................................................. 18  
   1.4. Core elements of Industry 4.0 .............................................................................................. 18  

2. **Process Mining** ....................................................................................................................... 20  
   2.1. Definition ................................................................................................................................. 20  
      2.1.1. Process modelling ............................................................................................................. 23  
      2.1.2. Data mining ..................................................................................................................... 25  
      2.1.2.1. Event logs .................................................................................................................... 26  
      2.1.3. Process Analysis .............................................................................................................. 27  
   2.2. Principles ................................................................................................................................. 28  
   2.3. Methodology ............................................................................................................................ 30  
   2.4. Challenges ............................................................................................................................... 31  

3. **Literature review** ...................................................................................................................... 33  
   3.1. acatech Industrie 4.0 Maturity Index ....................................................................................... 34  
      3.1.1. Introduction ....................................................................................................................... 34  
      3.1.2. acatech .............................................................................................................................. 34
3.1.3. Structure of the model.............................................................................................................. 35

3.1.4. Stages ........................................................................................................................................ 37
   3.1.4.1. Stage one. Computerisation ................................................................................................. 38
   3.1.4.2. Stage two. Connectivity ......................................................................................................... 38
   3.1.4.3. Stage three. Visibility ............................................................................................................ 38
   3.1.4.4. Stage four. Transparency ...................................................................................................... 39
   3.1.4.5. Stage five. Predictive capacity ............................................................................................... 40
   3.1.4.6. Stage six. Adaptability ......................................................................................................... 40

3.1.5. Capabilities .................................................................................................................................. 41
   3.1.5.1. Resources ............................................................................................................................. 42
   3.1.5.2. Information Systems ............................................................................................................. 43
   3.1.5.3. Organisational Structure ..................................................................................................... 44
   3.1.5.4. Culture .................................................................................................................................... 46

3.1.6. Functional areas ........................................................................................................................ 47

3.1.7. Application .................................................................................................................................. 48

3.2. Using the acatech Industrie 4.0 Maturity Index ............................................................................. 49
   3.2.1. Maturity analysis ...................................................................................................................... 49
   3.2.2. Areas requiring action .............................................................................................................. 50

4. Analysis. Usage of Process Mining within Industry 4.0 ................................................................. 51
   4.1. Industry 4.0 Maturity Stage limitations ..................................................................................... 51
     4.1.1. Digital Twin ............................................................................................................................ 51
     4.1.2. Cross-organizational scenarios ............................................................................................. 52
   4.2. Available solutions .................................................................................................................... 52

Conclusions and Outlook .................................................................................................................... 53

References ............................................................................................................................................... 54
Introduction

We are currently living in an era where the challenges confronted every day are increasing exponentially. Not only new businesses opportunities are emerging, but also traditional markets are being transformed. It becomes essential for any company to be able to quickly adapt itself and respond in a consistent way to the upcoming challenges to remain competitive.

The Fourth Industrial Revolution refers to the technological, social, economic and political phenomenon that has taken place since the First Industrial Revolution. Industry 4.0 would be the systematic, functional and practical implementation of the principles of the Fourth Industrial Revolution in the contemporary industry and its production, management and administration processes in order to achieve greater efficiency and instantaneity.

The developments related to the Fourth Industrial Revolution are not only intended to be achieved in a technological perspective. In fact, many companies have already successfully started individual digitalization process towards these aspects. Industry 4.0 also requires companies to transform their culture and organization within their all other functional and structural areas.

The development of the concepts of digital transformation and Industry to 4.0 is increasingly important for manufacturing companies, as they provide them with a clear advantage in the dynamic markets in which they compete. However, in practice it is difficult for companies to implement this concepts since Industriy 4.0 is more of a concept than a ready-to-deploy solution. In addition, its complexity makes it difficult to successfully implement Industry 4.0 systems that actually incorporate all organizational aspects and levels.

This makes it clear that there is a need to define a framework for guiding and analyzing Industry 4.0 implementations, which has to be based on an assessment of maturity, considering the capacity of each business.

The path to Industry 4.0 means that the physical and virtual worlds are going to start merging. More and more objects have intelligent technology thanks to sensors and actuators that allows them to be networked. In addition, not only are objects capable of leaving a fingerprint, but any event can be tracked digitally.
Once companies have achieved full digitization, absolutely all information will be available in real time from the cloud. It seems clear to question then what information is really relevant, and whether it is possible to give a meaning to all of this abusive amount of information.

In the face of this challenge, technologies are required to work with large amounts of data and to assist decision-making with a global application in any area of a company. After all, at the corporate level, it is the strategic decisions of great impact that will make the company a fully adaptable entity thanks to its hyper connectivity.

In this context, Process Mining becomes a perfect candidate, providing what is needed and working with the available resources.

**Motivation**

Industry 4.0 started to be presented at technological academical studies a long time ago, and it might not look like a contemporaneous topic anymore. Its scope is too broad, and sometimes it becomes easy to get lost between all of its features. Despite of this, it is not something that have been yet fully achieved in real companies and there is a need of research on this topic.

I personally find motivating to contribute to the state of the art of technologies that are not only making struggle corporations in the present, but that will shape the future.

I have had the chance to participate in a couple of professional experiences in which I have realized how Industry 4.0 can really make a change and not only improve, but skyrocket the whole business process.

Among this, Process Mining is a really young technology, that sits in a hype phase while starting to be implemented in some of the top companies worldwide. I find fascinating to be able to academically work with in current trend that promises to be crucial for the technological revolution.

**Problem statement**

Industry 4.0 concepts have been around for decades, but it has not been until a couple of years that they were all categorized under the name of the fourth industrial revolution we are currently living. Within this revolution, there are many technologies available and it becomes difficult to set an start and an end to any development process through it.
Companies are not aware about the exact roadmap, and there is a need of providing a modular system for applying a basic Industry 4.0 value creation design model.

There have been already studies on this, which build a framework upon which to keep researching. In fact, using this models, it could be seen how companies are behaving through this path.

Actually, this is the main problem that this projects aims to address: where are companies struggling the most in their path to Industry 4.0 and how could they do the next step.

**Objectives**

The main objective of this project is to analyse what is the current situation regarding the implementation of Industry 4.0 and study how Process Mining technologies can help to achieve it. The purpose of this study is to open a debate and discuss why companies that look forward to become Industry 4.0 standards might not be benefiting from technologies such as Process Mining, which can facilitate their journey.

The following sorted series of goals will lead to the accomplishment of the main objective:

- Study the current Industry 4.0 scenario.
- Present an adequate framework in which to classify the journey of companies towards Industry 4.0.
- Analyse how companies are behaving through this journey.
- Identify the main challenges that are being faced.
- Study the usability of Process Mining technologies.
- Define how Process Mining can help go through those challenges.
1. Industry 4.0

1.1. Definition

Industry 4.0 is the promised result of a new revolution that combines advanced production and operations techniques with intelligent technologies that will integrate into organizations, people and assets.

This revolution is marked by the emergence of new technologies such as robotics, analytics, artificial intelligence, cognitive technologies, nanotechnology and the Internet of Things (IoT), among others. Organizations must identify the technologies that best meet their needs to invest in them. If companies do not understand the changes and opportunities that Industry 4.0 brings with them, they risk losing market share.

For traditional leaders, accustomed to data and linear communications, the change in this new industrial revolution, providing real-time access to data and business intelligence, will transform the way they conduct their business. The digital integration of information from different sources and locations allows business to be carried out in a continuous cycle. Throughout this cycle, real-time access to information is driven by the continuous and cyclical flow of information and actions between the physical and digital worlds. This flow takes place through a series of iterative steps known as PDP (physical-to-digital-to-physical) as seen in Figure 1.¹

1. From the physical to the digital world. Information from the physical world is captured and a digital record is created.

2. From digital to digital. In this step, information is shared and interpreted using advanced analytics, scenario analysis, and artificial intelligence to discover relevant information.

3. From the digital world to the physical world. Algorithms are applied to translate digital world decisions into effective data, stimulating actions and changes in the physical world.

It is important to understand the potential of this fourth industrial revolution because it will not only affect manufacturing processes. Its scope is much broader, affecting all industries and sectors and even society. Industry 4.0 can improve business operations and revenue growth, transform products, supply chain, and customer expectations. This revolution is likely to change the way we do things, but it could also affect how customers interact with them and the experiences they expect to have while interacting with businesses. Beyond that, it could lead to changes in the workforce, which would require new capabilities and roles.²

In addition, technologies related to Industry 4.0 can also lead to completely new products and services. The use of sensors and portable devices, analysis and robotics, among others, will allow improvements to products in a variety of ways, from prototyping and testing to the

---

incorporation of connectivity to previously disconnected products. These changes in products translate, in turn, into changes in the supply chain and, consequently, to customers.  

The impacts of Industry 4.0 can be felt at multiple levels: in large ecosystems, at the organizational level and at the individual level (in employees and customers):

- **Ecosystems.** In addition to the change in which companies operate and in the production of goods, Industry 4.0 affects all ecosystem agents (suppliers, customers, regulatory considerations, investors, third parties...). These technologies allow interactions between each point in a network.

- **Organizations.** The ability to adjust and learn from data in real time can make organizations more responsive, proactive, and predictive. It also allows the organization to reduce its productivity risks.

- **Individuals.** Industry 4.0 can mean different things to each other. For example, for employees it can mean a change in the work they're going to do, while for customers it would mean greater customization in the products and services that best meet their needs.

### 1.1.1. Benefits

Taking into consideration that the current ecosystem in which companies perform is becoming completely unpredictable, the main goal to achieve with Industry 4.0 is to allow a company to be flexible and agile to continuously learn from the environment and adapt to any unexpected situation. Hence, adaptability is the main benefit of Industry 4.0. As it can be seen in Figure 2, the more time it takes for a business to react to a change, the more value losses it carries. Thanks to the implementation of Industry 4.0, any latency will be shortened.

---


11
With this been said, there are other key points that can be highlight as beneficial for an organization:

- **Meeting individual customer requirements.**
  
  To start with, we can consider as a key point the capacity to meet individual customer requirements. By implementing Industry 4.0 in the company, it widens the perspective in which the company can face a customer, as specific requirements can be included in the planning, manufacture or design stage, for instance, and is capable of manufacturing low volumes of production.

- **Flexibility.**
  
  Flexibility is also one of the essential benefits from which companies can gain profit. Creating a networking adapted to its needs, allows to set up different parts of the company or its processes depending on the situation.

- **Optimised decision-taking.**
  
  Another factor to bear in mind is the optimisation that is gained for decision-making. To determine and study which will be the best decision to take is one of the essential demands if you want to evolve in the global market. What Industry 4.0 provides to the company is a constant transparency which helps to make early-stage decisions with a fast response.
• Resource productivity and efficiency.

Regarding the direct benefit that is demanded from the companies, Industry 4.0 allows to perform the highest output of the products from the volume of resources available. Always considering that the company will use the lowest amount of resources to generate this output. In this way, manufacturing processes are being optimized in the most efficient way across the organization.

• Creating value opportunities through new services.

Industry 4.0 not only allows to create new value opportunities regarding the product or service, but also across the organisation.

• Responding to demographic change in the workplace.

It also enables to find solutions when there is a demographic change regarding the company. It is capable of turning demographic change, many times considered as a challenge, to a positive income for the company, because both the technical systems and the collaboration with people throughout the organisation will make it possible to create new initiatives.

• Work-Life-Balance.

It will also help to create a work-life balance as the organisation will be able to be more flexible allowing a better balance between work and private life. Also, it will allow to develop and motivate personal development throughout the employees.

1.1.2. Challenges

The implementation of Industry 4.0 embraces a broad range of technologies and procedures, some of which are still in a developing phase. Some of the most common challenges that companies and organizations have faced during this process are listed below:

• The development of software and analysis systems that allow the conversion of the immensity of data produced by smart factories into useful and valuable information.

• Respond to the current problem considering energy savings and an efficient natural resource management.

• Cybersecurity issues due to the need for industries to introduce open source systems to adapt to the digital transformation of their processes.
• Avoid computer or communication failures due to the large immersion of software and IT infrastructure to which companies' IT systems have to undergo.

• The lack of qualified personnel to lead companies towards the fourth Industrial Revolution.

• Reluctance of company managers to embark on the path to digital transformation of businesses.

• Certain companies (especially SMEs) must start by optimizing their processes or adapting their infrastructures and/or production facilities (Industry 3.0) first, before embarking on the path to Industry 4.0.

• Massive loss of jobs due to process automation where there will have to be a change in the tasks performed today by the manufacturing line operator. The vast majority will have to convert professionally, adapting to new jobs that do not exist today and that will arise with the paradigm shift.

• Ensure the return on investment (ROI) generated by the new technology required for the digital transformation of companies. Industry 4.0 requires high investments and usually provide long-term benefits.

In addition to the points mentioned above, it should be noted that today the vast majority of companies do not have state-of-the-art infrastructures in technology and, as a result, they will have more problems and will require more investment to achieve the digital transformation of their production processes. This revolution requires state-of-the-art industrial automation systems that can integrate more and more sensors, new technologies, and that have wireless communication capability. Factories must gain in interoperability capacity and massively collect data from the elements that make up their production processes to achieve real improvements in manufacturing efficiency and flexibility, and in turn must be able to manage and analyze these large amounts of data.

1.1.2.1. Industry 4.0 maturity

Among all of the challenges already mentioned, there is a specific one that may sit above any of them. Many companies are confident about starting the changes needed to make their facilities mime the Industry 4.0 standards, because they are looking for the attributed benefits already known or simply because they feel obligated to by the current market scenario.
However, the road towards Industry 4.0 is not clearly defined. The goal of organizations could be common: to become a hyperconnected automatically adaptable company. But their current stage within this transformation and the exact changes they need to execute will be different.

During this path, many challenges as the ones described will be faced. Hence, there is the need to plan beforehand how the approach is going to be like and how this evolution towards Industry 4.0 is going to be measured.

Besides specific individual projects about the implementation of one or some Industry 4.0 technologies or methodologies, a global approach needs to be described.

1.2. Industrial Revolutions in history

An Industrial Revolution is known as a historical period in which an uncontrollable and innumerable amount of technological, cultural and socioeconomic transformations takes place for the first time. These transformations had never taken place before and hence determine a transition in history.

The new concept of Industry 4.0 associated with the Fourth Industrial Revolution and cyber-physical systems is expected to be able to drive fundamental changes at the same level as the previous three revolutions. Now, it becomes interesting to take a look at the past and study how the previous industrial revolutions took place and shaped our present, as can be seen in a timeline in Figure 3.
1.2.1. First Industrial Revolution

The First Industrial Revolution (known as “Industrial Revolution”) drove the humanity from a rural world to an industrial world. It was a time of great inventions, highlighting one in particular, the steam machine. With this invention, the railway and the steamship were born. The social and economic change that took place at this stage was radical the emergence of a new social class, the proletariat, of a new conception where the world is getting smaller.

1.2.2. Second Industrial Revolution

The Second Industrial Revolution was characterized by the improvement of production thanks to the incorporation of new energy sources such as electricity and oil, the obtaining of new materials such as steel and the application of new forms of work such as automation and chain work. In addition, a new type of capitalism emerged called financial capitalism. This was the result of large investments made with the profits brought by the First Industrial Revolution and money from banking.
1.2.3. Third Industrial Revolution

The Third Industrial Revolution was based on new information and communication technologies, as well as on innovations that allowed the development of renewable energies. As a consequence, the potentialities of these two elements acting together, great changes are foreseen in various areas. Never before have such high levels of interactivity and intercommunication been reached, while innovations in energy could mean such a substantial change as is envisaged with the development and exploitation of renewable energy sources.

1.2.4. Fourth Industrial Revolution

The so called Fourth Industrial Revolution is characterized by a combination of technologies that is blurring the lines between the physical and digital domains. Its foundations come from the Third Industrial Revolution, that started a digital revolution around fifty years ago. Despite of this, the Fourth Industrial Revolution does not merely represent an extension of the previous one, but a distinct one due to its scope, the impact its having in the current systems and its speed of current innovations that has no historical precedent.4

1.3. Industry 4.0 around the world

The broad congregation of technologies and methodologies used towards the evolution to smart companies beyond digitalisation, have been lately named under the label Industry 4.0. Before the designation of this topic, some of its practices were already present, which means that there was a transition from the individual or singular applications to a certain intensity of development in the field. This led to name this stage a revolution and categorize it after the three previous ones in the history of humanity.

In this context, this revolution started to advance until being established in specific different ways, depending on the location around the world. It can be said that there is a common goal shared among all of them, but each territory shaped and named the revolution accordingly. Some of the new perspectives of the Industry 4.0 topic emerged worldwide: “Smart Manufacturing” in United States, “Innovation 25” in Japan, “Made in China 2025” in China, “Manufacturing Innovation 2.0” in Korea and “Industrie du Futur” in France.

Although there are some slightly different theories while approaching this topic, all of them take into consideration the same core elements and use similar definitions for them.

### 1.3.1. German industry

The German government, together with organizations that served as advisors, was the responsible of naming and defining this revolution back in 2011, calling it “Industrie 4.0”.

The Federal Ministry of Economy and Energy and Federal Ministry of Education and Research of the German government, together with high qualified representatives from companies, science/technological associations and trade unions, formed the Platform Industrie 4.0 to shape the future of the Fourth Industrial Revolution by developing strategies, technical solutions and recommendations towards the implementation of the Industry 4.0.5

The Industrie 4.0 Platform research advisory board has positively contributed to the state of the art of the implementation of Industry 4.0 methodologies. Thanks to the partnership with German leading industrial corporations and the implication with the public administration, its publications with strategic recommendations are having an important impact on not only Germany, rather the whole world.

### 1.4. Core elements of Industry 4.0

Industry 4.0 encompasses different dimensions, but in the technological field, it stands upon various pillars. These are different innovative technologies that become available in the Fourth Industrial Revolution:6

- Internet of things.

IoT describes the network of physical objects that are integrated with sensors, software, and other technologies in order to connect and exchange data with other devices and systems over the Internet using standard communication protocols. The use of these devices implies a change in the management and processing of data and information, formerly done by humans and now through the network of interconnected intelligent objects itself.

---

• Big data analytics.

Big data refers to datasets or combinations of datasets whose size (large volume of data), complexity (different sources and formats with greater granularity) and instability (high speed of generation and update) make it difficult to capture, manage, process or analyze using conventional technologies and tools. This data is valuable for improving decision-making and optimizing processes in business, so new techniques and technologies are required to analyze it.

• Advanced robotics and cobots.

A robot is a programmable automatic machine capable of performing certain operations autonomously and replacing humans in some tasks, especially heavy, repetitive or dangerous ones. Generally, it is equipped with sensors, which allow it to adapt to new situations. Two types of robots are often differentiated: industrial and service robots. Industrial robots are used in industrial automation applications in handling or manufacturing processes and are often in the form of multi-axis multipurpose manipulators. Service robots, are used in complex environments interacting with dynamic people or objects and this forces them to have some autonomy. On the other hand, cobots (collaborative robots) are robots designed to have direct contact with humans. They provide great adaptability and their applications differ completely with those of the traditional industrial robot.

• Machine learning

Machine Learning is a scientific discipline in the field of Artificial Intelligence that creates systems that learn automatically. Learning in this context means identifying complex patterns in millions of data (big data). The machine that actually learns is an algorithm that reviews the data and is able to predict future behaviors. Automatically, also in this context, it implies that these systems are improved autonomously over time, without human intervention.

• Cloud computing

Cloud computing enables users to remotely use and access computer hardware and software over the internet. Instead of having to purchase, install, run, and maintain various computing resources locally, the user can contract such cloud services to a
provider that owns and maintains them. In this way the user can access a complete and specialized resource source according to their demand.

- **Artificial intelligence**

Artificial Intelligence (AI) is the combination of algorithms posed for the purpose of creating machines that present the same capacities as human beings to reason and make decisions. It encompasses a set of software, logic, computing and philosophy disciplines to understand human cognitive abilities in order to replicate or improve them.

- **Additive manufacturing (3D printing).**

Additive manufacturing is an industrial production concept through which the material is deposited layer by layer in a controlled manner where it is required to create three-dimensional objects. Using this technique, which we know as 3D printing, you can produce geometric shapes customized according to the needs of each sector. Compared to traditional manufacturing techniques, this technology reduces intermediate processes such as tooling production and allows only the material to be used for the manufacture of the part without generating waste.

- **Computer-aided design (CAD) and Computer-aided manufacturing (CAM).**

CAD/CAM software is used to design and manufacture prototypes, finished products, and production runs. An integrated CAD/CAM system offers a complete solution from design to manufacturing. CAD consists of a series of software solutions that make it easy to design parts or structures and even simulate tests of various types. The CAM consists of the software that transfers the CAD design to the machine tool responsible for producing it, indicating all the corresponding parameters.

## 2. Process Mining

### 2.1. Definition

Process Mining is a reasonably new family of techniques that lies between two already well known research disciplines: process science and data science. On one hand, there are process modelling and analysis with quite traditional masteries and formal methods, whereas on the
other hand, there are modern and innovative branches based in computational intelligence and data mining.⁷

![Diagram of Process Mining foundations. Process Mining. Data Science in action.](image)

The purpose of Process Mining techniques is to discover, improve and monitor processes not only in businesses, but in any existing application area. With the current digitalization development it becomes possible to automatically extract information from available technology systems. Process Mining uses this information and transforms it into knowledge.⁸

The interest in Process Mining have been increasing since its academic origins to the creation of companies offering this technology as a solution to well established corporates. Two main drivers have facilitated this natural evolution. On one hand, since there are events constantly being recorded in companies, building a historical data framework about process suitable for Process Mining techniques implementation. On the other hand, due to the rapidly everchanging

---


environments and the growing market competition, companies ask for support in their business processes.\(^9\)

Process mining sets its origin from the difficulty to have optimal processes. Assumed processes that businesses use to analyse their performance across their organization units are too idealized from the real ones and do not consider all existing variations. These processes do not exactly represent the real world and any study made with them would imply not understanding the reality.\(^10\)

Process visibility has been hampered by building processes on top of rigid and fragmented transactional systems. A typical process, may touch any number of IT systems, different users and different departments. These makes it really difficult to get an overall visibility of what is currently happening in the business processes. This is ultimately the problem that process mining aims to address.

With Process Mining it can be looked into those transactional systems and measure the capacity of the processes to see how processes really run. Hence, gaps can be identified to know which ones have the greatest impacts in the core business operations. Therefore, action could be taken in real time to unlock execution capacity in the processes.

Before process mining, the traditional approach could be to do a process mapping exercise. It involved meeting with stakeholders to bring together information about a process and build a model. These method relies on people, whose insights might be subjective and partial, and it is time consuming. In addition, it delivers a one-time only understanding. It does not adapt to the changing process since another meeting is required to remap a process.

In contrast to this, process mining relies on the data produced by the process to create an objective and complete view of the business operation. Since it is based on data that can be exported automatically in real time, its results are immediate. As the process changes, the analysis changes and its outcomes too. This means that once the model is setup, it can continuously be used enhance the process.

The basic Process Mining technique relies in four simple steps represented in Figure 5. Within a company, any activity that is performed generates data that can be tracked. This data is

---


categorized in terms of event logs, related to each activity. From these data, a process model is built and then analysed to bring insights and drive decision making to improve the process performance.

In the following three sections the three pillars of Process Mining are presented. They include process modelling, data mining and process analysis.

### 2.1.1. Process modelling

Under the scope of production and operations management, a process model is a representation of series of linked actions or steps taken in order to achieve a particular end.

There are actually many processes in our own daily lives. Processes are absolutely critical in our day to day lives and exactly the same is true for modern day enterprises. Processes are fundamental to enterprises execution.

The operations research and management is branch of the business management sciences that comprehensively relies on modelling. Models are used to study current processes within a business and as a support to make decisions inside processes while performing planning or
controlling tasks. Then, using simulation, a process model can be used to analyse its performance beforehand and estimate or predict how it behaves in front of certain fluctuations.\footnote{Van der Aalst, W. (2011). \textit{Process Mining: Discovery, Conformance and Enhancement of Business Processes}. Springer Verlag. Chapter 2}

In first instance, processes are implemented to get intrinsic benefits from repetitive actions. Under a manufacturing company point of view, process improvement happens towards three main objectives: to reduce the time it takes to deliver a product or service, to reduce the cost of delivering those goods to a client and to improve its quality. Organizations go about improving processes through changing the processes themselves. Actions like standardizing, streamlining, optimizing or automating are performed to do so.

The most common representation of a process model is a transition state flow in which using a specific notation (commonly arrows, connectors, operators, activities and terminators) a business process can be illustrated. A simple example can be seen in Figure 6.

![Sample process](image)

\textbf{Figure 6. Sample process. Process Mining: Discovery, Conformance and Enhancement of Business Processes.}

In the real world, processes tend to be complex as many variables interact and personal interference might be sometimes hard to model or even unpredictable. In addition, in some cases the use of a model might be inadequate since it could be describing an idealized version of the real process.

Instead of starting from a model, process mining suggests to extract the models needed based on the information from event logs. The combination of all the event logs cover all the cases that ever occurred in process, so any model can be build based on these facts.
Model analysis relies on the availability of good quality models verified to be aligned with reality. In case a model does not properly represent the reality, it becomes useless as any performance analysis will not be significant.

Since process mining stands for building a model based on the data from the lowest possible level, it establish a direct link between the model and the real process. Through filtering the event logs, the process can be studied from different points of view by getting a model for a certain scenario.\(^\text{12}\)

### 2.1.2. Data mining

The growth of the digital world has driven to the existence of massive amounts of data that are impossible to study with conventional analysis methodologies. Data mining is the process of discovering patterns, anomalies and correlations in large data sets through specific methods to understand and predict outcomes. This field of study sits between statistics and computer science disciplines making use of machine learning and other intelligent methods to work with data.\(^\text{13}\)

Data mining techniques filter chaotic disorganized amounts of data to comprehend what is relevant and assess further use such as decision making.

A data set consists on a sample of instances such as objects, cases, records or entities that have attributes known as variables and are usually presented in a table. In case the data is labelled, supervised learning techniques can be applied to explain a response variable in terms of other predictor variables. In case the data is unlabelled, this means that variables cannot be split and unsupervised learning techniques should be applied.\(^\text{14}\)

Some of the methods or techniques used in data mining are clustering, affinity grouping, regression, neural networks, decision trees and predictive analytics.

Whereas data mining is a data-centric procedure, process mining focuses on a process-centric approach. Undoubtedly, process mining techniques can benefit from knowledges from the data mining methodologies and apply them to end-to-end processes.

---


\(^{14}\) SAS (2021) *Data Mining. What it is & why it matters*. SAS Institute Inc.
2.1.2.1. Event logs

For every single lowest level activity that happens within a business unit, an event log is generated. In order to apply process mining practices it becomes indispensable to record and save those event logs from the different data sources available.\(^\text{15}\)

In standard companies, event data can be found everywhere. Event logs can be extracted from a database system, a spreadsheet, a transaction log, or basically any business suite commercial software such as ERPs or CRMs. Once the relevant data is available and has been found, the export and conversion is rather simple since there are already some standards recognized for data formats. The challenge relies on selecting the event data related to the process of study. A data flow overview can be seen in Figure 7.

An event log needs to contain at least a minimum requirements which comprise having an identifier, a correspondent timestamp and an activity name. The identifier or case id could be an order number or an invoice number. An event’s activity could represent, for example, a start, resume, complete, skip, suspend, abort or any other activity related information. The timestamp characterises when the activity occurred.

The three fundamental elements mentioned are required in the data in order to conduct process mining, but in addition, an event log may contain many other optional variables. There is also

the case that through cross references more data can be related to an event log. Any meta data related to the process will be then combined to build a single process data model.

An example of an event log has been represented as a table in Figure 8. The event log derives from a process which consists of cases. For each case, different unique events occurred and these are characterized with its own attributes (identifier, timestamp, activity and other optional properties).

![Figure 8. Event log example. http://www.processmining.org.](http://www.processmining.org)

The set of event logs determine all of the possible variations of the process. These different variations can be aggregated in a superimposed process map [X] which includes a track of absolutely all the activities. These model can be then sliced and analysed to understand the impact of each of its characteristics.

### 2.1.3. Process Analysis

With a process visualization, we can see the chronological sequence of events as it is actually happening in the real business. By using this meta data, an in depth analysis can be performed on the process.
Process analysis and improvement can be done through different approaches. Basically, there are three types of process analysis that apply to Process Mining and they are represented in terms of input and output in Figure 9.

![Process Analysis Methodologies Diagram]

Figure 9. Process analysis methodologies. Process Mining Manifesto

One approach would be a process discovery examination. For every case, there is a track record that shows at which time was every different activity through the process done. This allows to study the process in terms of time and determine if there are cases running late, for example. In this way, it could be discovered the existence of bottlenecks or if any resource seems overloaded.

Another approach would be to conduct a process conformance checking in order to know if a process is performing as expected. For example, it could be tracked which activities are often skipped or which resources are causing a deviation.

Through these insights, process mining aims to give recommendations on where to enhance the process. These recommendations can involve minor changes such as adjusting or automating certain activities, or they could even involve a redesign of the whole process. In this case, an optimal model would be generated.

### 2.2. Principles

Since Process Mining is a really innovative technology, there exists some fundamentals that serve as a foundation when implementing it. These good practices are exposed next as a list of
guiding principles that should be followed to grant reliability on the usage of Process Mining techniques and to avoid already known mistakes.\textsuperscript{16}

Stage 1: Event data quality should be audited.

Event logs collection should be intentionally designed since the validity of the process relies on its quality (and not quantity). Their correctness and accuracy must be assured.

Stage 2: Data extraction should be goal-oriented.

Before using any Process Mining technique, its utility should be planned in order to give answers to different concerns about a process. Then, the specific type of cases to be analysed could be selected.

Stage 3: Basic control-flow operators should be supported.

While defining a process, main languages should be considered so the model is compatible and understandable. Besides main logic trees and gateways, there are specific elements in complex processes that need to be defined properly.

Stage 4: Event correlation should be rigid.

There must be a relationship established between the elements in the model and the events in the log in order to work with a comprehensive representation of the process and facilitate analysis.

Stage 5: Models should be treated as representations of reality.

Given a range of event logs, many valid views of the process might be useful, but all of them will be representations of the reality since they derive from event data.

Stage 6: Process Mining should be a continuous process.

Since Process Mining is based on event logs, and those are constantly being generated and collected, the models should be continuously updated. A process might change anytime due to its dynamic nature.

2.3. Methodology

Process Mining procedures follow a life cycle, and its essential stages can be seen in Figure 10, showing also the concrete activities related to each of them.

Respecting the principles already defined, every Process Mining project initiates with a plan (Stage 0) in which the objectives and the boundaries of the project are defined. After having started the project, the first step is to extract the data (Stage 1) from the current processes. This involves understanding the available data and keeping in mind the objectives of the project. The results of this stage will be a historical dataset, examples of handmade models, objectives, and questions. The next step is to create a control-flow model and connect it to the event logs (Stage
2). Process analysis techniques may be used in this stage such as the ones related to process
discovery, in which the model discovered itself could already provide some valuable
information. In this, a link between the model and the event logs is built, to correlate events to
real process activities. Depending on the complexity of the control-flow model, there is the
possibility to extend it with other perspectives and create one or various integrated process
models (Stage 3). These might be built with filtered event logs or aggregated data to get a
specific model that give answers to specific problems stated in the plan. Both models generated
in Stage 2 and Stage 3 might trigger some actions such as adjustments or even a redesign of the
process that might drive to loop back to previous stages. Once the project has gone through
these stages in a sufficiently structured and stable manner, the next step is to provide operational
support (Stage 4) by performing analysis to get predictions with the current data.
Recommendations given in this stage will cause an intervention on the process to improve the
real process.

Through this Process Mining methodology any real process will be first discovered, then
monitor and finally improved

2.4. Challenges

Process Mining is an emerging technology that many companies are already benefiting of.
Because of digitalization, there is being a growth in the availability of event data across many
industries. This requires companies that work with Process Mining to define reliable procedures
so that they build a framework upon which to applicate updates and keep working
consistently.¹⁷

In this context, many challenges are rising, and some of them have been collected in the
following bullet list.

- Finding, merging, and cleaning event data.

The extraction of event data and the following categorization of event logs is the
foundation of Process Mining and it usually presents some obstacles to overcome. Data
is usually distributed over different sources and compatibility issues are faced when
merging it. Event data tend to be object-centric and not process-centric, which can lead

¹⁷ acatech, Plattform Industrie 4.0 (2019). Key themes of Industrie 4.0. Research Council of the Plattform Industrie
4.0
to complex event logs with diverse characteristics difficult to analyse. Some event logs might even be incomplete.

- **Benchmarks availability**

  There is not yet a broad state of the art regarding Process Mining techniques where solutions from different vendors could be easily compared. There needs to be a consensus about the quality of these techniques.

- **Understanding the drift.**

  Processes are never in steady state by nature. This introduces the concept of drift, which means a process changes while it is being studied. In this context, other changes may appear such as seasonal variations. This should be managed through high quality event logs and integrated process models.

- **Process representation bias.**

  The process model obtained from the event logs using a particular process discovery technic might be differed from the visualization of its representation. Depending on the language used and its compatibilities, a graphical representation might be wrong.

These list represents some of the current challenges that are currently being faced. Over time, many of these are expected to be solved, and others are there to come. The list does not intend to be complete, but to represent the main topics that need development.
3. Literature review

Extant research has extensively studied many topics related to Industry 4.0 and Process Mining in various contexts. However, academic research specifically focused on the available Process mining solutions and the maturity stage in Industry 4.0 required to implement them is currently scarce.

Whereas academic literature on Process Mining concrete solutions and how to implement them is currently scarce, practitioner literature contains a full descriptive work on the field. In particular, Industry 4.0 has been already described and analysed in both academical and professional corporative studies. On the other hand, Process Mining, as a more recent term, has been also defined and studied by many.

The literature review has consisted on four steps of searching for information.

First, in order to set a solid background about the field of study upon which to build the insights of this project, some specific related literature has been reviewed. These included popular studies written by high-profile practitioner experts that were relevant from the point of view of the Industry 4.0 state of the art. Many new concepts emerged during these first search, being of particular interest and arising among others the Digital Twin concept. The same sources of information served as an introduction to Process Mining, a term certainly related with the Industry 4.0 and the Digital Twin concept. Hence, many literature about Process Mining were also reviewed.

Secondly, specific contextualization about the Industry 4.0 was evaluated, focusing on acatech’s Industry 4.0 Maturity Index. These study has set the framework in which the statements from this project are based.

Then, using the literature already gathered, a citation searching technique was performed to deep into the main concepts related to this project.

Finally, and mainly guided by advertisements and online reviews, a general Google search was executed to discover specific applications of Process Mining within Industry 4.0 and find content related from companies that provided these services. Several companies’ websites were checked to understand their description of the topics from a business point of view and also to study the services that were providing.
3.1. acatech Industrie 4.0 Maturity Index

3.1.1. Introduction

The acatech Industry 4.0 Maturity Index sets a framework in development of businesses towards learning and agile organizations. It assists companies determining at which stage in the transformation towards a fully mature Industry 4.0 they are and it states the capabilities needed to progress within the index. Since Industry 4.0 involves a transformation, the index evaluates each subject from different perspectives such as organisational, technological, or cultural.

Although the acatech Industry 4.0 Maturity Index intends to provide a general model collecting all the meaningful concepts, any company's journey to Industry 4.0 will be different. As a result, it becomes essential for each company to start by studying its current situation and its strategic objectives. These will define the areas in which the evolution towards Industry 4.0 will add value.

The next step after identifying the objectives is to analyze what state of maturity the company is considering how the various systems and technologies related to Industry 4.0 have been implemented in each area of the company.

Based on this analysis, you can identify the areas in which the company must develop to progress on the maturity scale and achieve its initial objectives.

3.1.2. acatech

The German Academy of Engineering Sciences, so called acatech (from the words academy and technology), is a non-profit organization whose main goal is to act as a competent and trustworthy consultant for politics and society to facilitate the necessary decision-making in the technological challenges that society faces. The association brings science, economy and society together to work on projects towards a sustainable growth and prosperity not only nationally, but also internationally.\(^\text{18}\)

The academy is made up of outstanding scientists of recognized reputation for their scientific achievements. The members of the academy belong to the engineering, natural sciences, humanities and social sciences branches. However, professionals from technology-oriented companies and scientific associations or organizations also collaborate.

\(^{18}\) acatech - Deutsche Akademie der Technikwissenschaften (2021).
Since the academy sets its consulting mission independently, its projects outcomes can be considered to be oriented to the public interest. This is a key fact in terms of educational purposes since there are no private concerns and all the knowledge is published transparently. In fact, the academy publicly publish articles and papers about the projects that are being developed. It has been really helpful for this project to have access to the “acatech Industry 4.0 Maturity Index” study publication as well as its case study publication.

Regarding what concerns to Industry 4.0, acatech has partnered with Platform Industrie 4.0 to provide strategy advisory for its implementation.

### 3.1.3. Structure of the model

The model is based on a series of maturity stages, these stages help the organisation create its own roadmap to determine how it will develop its transformation. Every company must state which are their objectives depending on their business strategy and, also, to consider which maturity stage represents the best stability between: costs, capabilities and benefits.

In addition, the model follows a structure based on the “Production and Management Framework”. This Framework is based in four structural areas that allow to analyse and determine the guiding principles that will provide the company with the Industrie 4.0 capabilities that it has still not developed.

The acatech Industrie 4.0 Maturity Index is based on the “Production and Management Framework”, which splits the internal parts of a company into: corporate structure, corporate processes and corporate development. It also considers the stakeholders groups and the environments where the company performs. Nevertheless, it is important to say that these last points are not considered in the Maturity Index scope and, therefore, are not being included in the present study.

When talking about corporate structure, this refers to those parts of the company considered to be essentials to produce the product or services. Whereas, for instance, corporate processes refer to all the processes chains that build up the company. On the other hand, corporate development refers to the strategy, optimisation, operations and rearrangement. Corporate development’s role is to be able to maintain a solid level of performance and in a continuous way, being capable of changing and adapting if needed.
When considering the Industrie 4.0 Maturity Index, it is in its basis to embrace a digital transformation strategy and implement it through the corporate development process of the organisation to achieve the optimisation and rearrangement factors.

The acatech Industrie 4.0 Maturity Index splits the company’s structure into four essential structural areas: resources, information systems, culture and organisational structure. For each structural area, there have been identified the capabilities that are needed and two principles. These capabilities support the companies with the starting point to transform the organization.

In addition, the corporate processes are the core of the five functional areas that are being studied by the Index. The five functional areas are: production, logistics, services and marketing and sales.

Figure 11. Model overview. acatech
When defining the implementation strategy, both structural and functional areas, as well as the maturity stages are merged to study in which level of maturity is the company in and, also, to consider the maturity level of each stage.

This model allows a graphical representation of the current maturity stage of the company using a circular diagram as shown in Figure 11. Each sector of the circle represents a structural area, and the radius within each sector sets the maturity stage of the particular structural area. Each of the axis of the sector defines a principle. The capabilities within that principle must be achieved to gain maturity in Industry 4.0.

3.1.4. Stages

The roadmap that has been defined by the model follows a six-step stage process. Each of these stages is linked to the previous one and created based on the results gained in the previous step. The step must describe the capabilities that are needed to achieve it and which will be the benefits for the company.

In the Figure 12 there is an overview of all the maturity stages. As it is stated, Digitalisation does not remain part of the Industry 4.0, it is considered to be part of the Industry 3.0 and both Computerisation and Connectivity are previous steps that are essential to start with its implementation. After these two steps have been completed, they will be followed by four differentiated stages that constitute the capabilities needed for Industry 4.0.
3.1.4.1. Stage one. Computerisation

Computerisation remains the first stage in the development path, providing the company with the basic needs for digitalisation. To proceed with the stage, there are different information technologies that are used separately from each other in the company. As it is logic, computerisation is already a step that has already been applied in many companies and is normally used to carry out repetitive tasks in a more efficient way.

At this stage, companies tend to have a traditional organisational structure dedicated to develop an efficient operation of each department individually.

The outcome that must result from this stage is a support from data processing systems by which the employees are liberated from repetitive, daily activities.

3.1.4.2. Stage two. Connectivity

The connectivity stage is used to replace the isolated deployment of information technology with connected components. This will allow that parts of the operational technology systems, known as OT, provide the company with connectivity and interoperability. In this stage, full integration of the IT and OT layers will have not occurred for now. It is important to state that this is considered an essential requirement for the Internet of Things. Also, it is important to mention that the Internet Protocol will allow to unify and standardise the communication.

At this point, the company’s status remains still traditional. The company still uses traditional project management tools that do not help make the process quicker and more agile. With that been said, it is also important to state that the company shows a will to incorporate change and move towards innovation.

In this case, the outcome that it is expected is that the IT systems have been structured and interconnected.

3.1.4.3. Stage three. Visibility

Following connectivity comes the visibility stage. Where the company has created a digital shadow from themselves. A copy that enables to know and see what is happening in the company at any time, providing real data that will help to carry out in a more precise way, management decisions. This digital shadow has been created with sensors that capture the processes form their start to end and also by recording events and states in real-time throughout
the company. It is also important to highlight that comprehensive capture data throughout the company is a key factor that allows provisioning of relevant data about the operation.

With this been said, companies must really change their perspective and way of thinking to embrace this stage. What has to be achieved is a capability of creating an up-to-date model of the company at every moment that must not depend on individual data analysis and avoiding, by any means, to only be capable of collecting data that only enables a particular analysis of an operation. We have to bear in mind that visibility does not only consist of a technological process, but it is also a process that shakes the organisational structure and the culture built during years by the company.

At this point, the company shows a current status where data can lead to various problems that must be addressed. From one side, we have to consider that the sources providing data can be several and different as data is usually stored in decentralised silos. Also, departments such as logistics, services or production can often lack of data been collected and the data collected is normally accessible for a limited number of people creating a difficulty for data to be used.

On the other side, the main outcome resulting from this stage is that the company acquires a digital shadow which enables strategies and decisions to be carried out with updated data.

3.1.4. Stage four. Transparency

Transparency is a stage used for the companies to understand how and why something is happening and, consequently, learn how to use this information that will be provided by analyses. The data presented in the digital shadow has to be analysed by means of engineering knowledge based on semantic linking and aggregation of data that provide information about the process that is required to make a quick decision-making.

That been said, the new technologies that help support the analysis of big volumes of data are helpful in this stage as they can be useful to find interaction the digital shadow of the company and, therefore, create a state for transparency regarding these interactions.

Therefore, the company is at a point where they have to be willing to not only use big data for technological aspects, but also to use data at every stage or level of the company to carry out decisions. The company is been changed in every department and employee and that means that a collaboration between employees and departments is required.
The main outcome resulting is that the companies have achieved to understand and use knowledge to make their decisions.

### 3.1.4.5. Stage five. Predictive capacity

The predictive capacity is a stage where the company is now able to have perspective and simulate hypothetical different future scenarios to understand which one is the optimal. In order to do so, the company will have to future-project the digital shadow so that they can study and predict different scenarios, considered depending on how likely they are to become in the present. This provides the company with a capacity to anticipate future processes in order to limit negative impacts and provide efficiency.

It is important to consider that the capacity of a company to predict their future processes depends on the previous groundwork that has been carried out. This is why it is essential to have a solid, well-structured digital shadow that the company can also combine with the knowledge they acquire from relevant interactions. In this way, the company will be able to forecasts and make recommendations with a relevant data background.

The company status for this stage shows that the company starts to be able to predict their futured ant, therefore, minimize negative impacts or prepare with anticipation. For this to be implemented fully, the company must embrace change regarding their structure and culture, enabling employees to make decisions faster when a situation is presented.

The main outcome of this stage is the anticipation. The company is able to take decisions that are based on future hypothetical scenarios.

### 3.1.4.6. Stage six. Adaptability

Adaptability remains the capacity of the company to adapt to future events and innovation constantly in the most rapid and efficient way possible. We can consider that this stage has been undertaken by the company when it is able to use the data conveyed in the digital shadow to make optimal decisions that give the best possible solution in the shortest period of time, being able to implement these decisions in an automatic process. Adaptability can only be successful when the company collaborates in a dynamic way across their value chain.

In this last stage, we can consider that the level of adaptability acquired by the company will depend on how complex are the decisions that they have to take and which is the cost-benefit ratio. It is considered to be more beneficial to only automate individual processes and asses if
it is a risk to automate approvals and acknowledgments. As it has been stated in numerous occasions, embracing change has to be mandatory, creating cooperative communities and agile, efficient project management across the company. It is important to convey that the company and its employees will have to be in constant learning in order to create a knowledge-based culture.

With that been said, the outcome resulting of this last stage is the capability of the companies to respond autonomously to any event, having systems capable of self-regulating and operate in a solid way.

### 3.1.5. Capabilities

There are four structural areas that a manufacturing company must have to start the transformation process to become an agile organisation, capable of learning. These four structural areas are: Resources, Information Systems, Culture and Organisational Structure. These areas are connected with six stages of the development path followed by the Industry 4.0. Each of the structural areas has two principles that represent the capabilities that must be developed in each stage for it to be successful. Therefore, the level in which these capabilities are successfully implemented, will determine the maturity stage of each principle.

An overview of all the capabilities that will be studied can be seen in Figure 13.
3.1.5.1. Resources

Any tangible and physical resource is considered to be part of this structural area, such as human resources, materials, the product or machinery. The companies have to work to acquire an interface for their resources between the physical and digital dimension.

The two principles for maturity within this structural area and its capabilities are defined below and summarized in Figure 14.

- Digital capability. In order to support any decision-making, the resources must have capabilities that allow to work in a digital way, enabling to collect and process data.
  - Provide digital competencies. This is a key point where the company has to ensure that its employees have the appropriate IT skills to be capable of developing decisions based on data.
  - Automated data acquisition through sensors and actuators. To be able to generate any data feedback, the technological resources of the company should have a data processing layer added, that will, as a result, enable to create a digital shadow.
  - Decentralized pre-processing of sensor data. Such as the sensors and actuators, the embedded systems form a link between the communication layer and the electromechanical components and are also a key factor of the cyber-physical systems, allowing to decentralize the pre-processing. In this way, it is achieved a reduction of the signal propagation delay allowing that time-critical computation can be carried out faster.

- Structured communication. It is extremely important to achieve a communication media between the employees, the interfaces between employees and machines, as well as between the machines. This will enable to create temporary networks that allow these resources to interact with each other.
  - Efficient communication. The communication between different resources should be configured in order to achieve real-time data exchange.
  - Task based interface design. Both identification and visualisation technologies create the main interface between the real and digital world. To include a graphic interface will allow to provide context-sensitive information.
3.1.5.2. Information Systems

Information systems essentially prepare, process, store and transfer data and information so, the configuration of this system is critical to ensure that data is used to make decisions.

The two principles for maturity within this structural area and its capabilities are defined below and summarized in Figure 15.

- Self-learning information processing. Data should be captured, prepared and processed in a way that employees can use it to make decisions. That is why the company has to acquire the correct technical requirements for it to be possible.
  - Automated data analysis. To achieve a decision-making process based on data, the company must count with an automated data analysis that is continuously adding data to produce information. Therefore, this data streams must be monitored constantly using specific rules and correlations to ensure that it produces adequate data.
  - Deliver contextualised information. The information is prepared and delivered in a specific way meeting the requirements of the company which will allow to support the decision-making.
  - Task-specific user interfaces. The information must be permanently matched and adapted to the needs of whom is going to use it, always considering the employee’s skills.
  - Build a resilient IT infrastructure and implement situation-based data storage. A solid IT infrastructure must be continuously adapted to meet the needs of the company and to guarantee the system’s usability in a long-term period.
• Information system integration. The company’s system must integrate all the data to make a common use across the organization. Therefore, a central platform that connects the IT systems and the resources is necessary for agile companies.
  
  o Integrate information systems vertically and horizontally. Systems must be fully integrated in a vertical and horizontal way, creating a common and single source of information where all the users access.

  o Implement data governance. Data governance policies give the company a guidance to process, store, manage and present the data with high-quality standards.

  o Standardise data interfaces. IT systems have to be connected between them with standard data interfaces to ensure and agile information flow and data exchange.

  o Upgrade IT security. To avoid taking risks, having a compliance with standards like IEC 62443 can be of real use, because it contains measures to maintain the IT security or adapt it following changes.

Figure 15. Information systems’ capabilities. acatech.

3.1.5.3. Organisational Structure

Organisational structure will be required to allow the correct transformation of the company. This is referred to the company’s internal organisation and its position inside the value network. The two principles for maturity within this structural area and its capabilities are defined below and summarized in Figure 16.

• Organic internal organisation. In this type of organisation, the employees have more individual responsibilities and less constraints when working. For this organisation to be successful in the company, the employees have to be highly skilled and it is needed a dynamic environment.
o Flexible communities. In an agile company, change must be part of the culture of the company where tasks and the teams may be change on a regular basis. Therefore, the company has to work to facilitate communication.

o Decision rights management. Many decisions can be taken in a decentralised manner, but several of them, such as strategical ones, must be taken centrally.

o Motivational goal systems. Guaranteeing that employees remain focused and following clear objectives is needed to ensure the creation of value. There must be different kinds of goals to ensure that the employees improve continuously. Therefore, there must be a right balance between monetary and non-monetary incentives inside the company.

o Agile management. It is the application of the main factors of Agile software development into the management processes.

• Dynamic collaboration within the value network. An exchange of automated and seamless information between different companies can help to create a cooperation and increase market transparency.

o Focus on customer benefits. Every company must focus on analysing how can they increase the customer’s satisfaction and meet their needs. For it, companies must be capable of changing and adapting continuously.

o Cooperation within the network. To include systematically different partner’s companies in the value network of the company, allows it to respond more efficiently and rapidly to the market’s needs.

![Organisational structure capabilities](image16.png)

Figure 16. Organisational structure capabilities. acatech.
3.1.5.4. Culture

The company’s culture will play a key role to successfully adapt to change the mentality of its employees. The company must create a culture based on its employee’s skills and their trust for digital systems.

The two principles for maturity within this structural area and its capabilities are defined below and summarized in Figure 17.

- Willingness to change. To embrace change, it is important to have an open perspective and be capable of overseeing when change is needed.
  - Recognise the value of mistakes. Mistakes are the key factor that trigger change and understand better the company. For it to be done, the culture of the company must be open to discussion and determine the causes of the mistakes to find solutions.
  - Openness to innovation. It is always a requirement to build decisions facing the future and change to adapt the company to the market’s needs.
  - Data-based learning and decision-making. Employees are confident with the data and base in it to take decisions.
  - Continuous professional development. To constantly develop required skills for the employees to continue evolving and developing is mandatory when talking of the digitalisation of the industry.
  - Shaping change. The correct changes and implementations must be done to enable a fast and effective response to external change.

- Social collaboration. Building trust, relationships and collaboration between the organisation is the perfect ingredients to share knowledge and evolve.
  - Democratic leadership style. Equity when deciding and embracing a flexible approach for decisions have to be adopted.
  - Open communication. Sharing knowledge creates a direct and open communication atmosphere. Also, when having the correct skills and the information, employees can easily work together in diverse situations.
- Confidence in processes and information systems. Building confidence means that the employees understand and accept how the information system makes the decisions and recognize it as beneficial.

![Figure 17. Culture’s capabilities. acatech.](image)

### 3.1.6. Functional areas

A given company can be divided into five different functional areas: Development, Production, Logistics, Services and Marketing and Sales. Then, each principle has to be individually evaluated for each functional area, leading this to minor adaptations of the involved principles. These functional areas are cited as the main functional areas of companies, being transversal to the vast majority of them. In addition, depending on each firm, the industry in which it competes, the magnitude of its resources, the scale of its operation, the type of organization and other particular characteristics some functional areas can be not included in the core business, hence discarded.

![Figure 18. Functional areas of the model. acatech.](image)

- Development. It is becoming a common way of working, to consider not to have a rigid finished vision of the product, but to start working with an initial vision of it which will keeping evolving during its development.
• Production. Customer’s requirements must be the fact that conditions how the processes and manufactures will be developed.

• Logistics. Logistics is considered within the internal and external activity of the company. Its basic function does not change as it is to ensure the availability, place, time and adequate cost for the customer.

• Services. Adapting services to the requirements of the customers remains more effective when building relationships with the customers and delivering them.

• Marketing and sales. Its role will remain essential, as it will constantly guide the potential clients through a customer journey to achieve a purchase.

### 3.1.7. Application

Companies who undertake the process of developing towards Industry 4.0 may use the acatech Industrie 4.0 Maturity Index as a guideline by following the next application steps:

**Stage 1:** Identification of the current Industrie 4.0 maturity stage

The first stage requires to assess its functional area and identify the maturity stage in which is currently. For it to be done, the questionnaire may be used to rate the capabilities of the processes arriving to a final score for each process and for the company.

**Stage 2:** Identification of capabilities requiring development

It consists of the company identifying which is the target development stage that wants to achieve at the end of its transformation. This requires analysing its current Industrie 4.0 capabilities for each functional and structural area and developing a gap analysis to study which are the capabilities that are missing, with the objective of developing them in a similar rate.

**Stage 3:** Identifying concrete measures

This stage highlights the actions that must be taken and incorporates them into a roadmap to start developing the process. After this, the measures that have to be taken in each area must be determined and, consequently, start implementing them.
3.2. Using the acatech Industrie 4.0 Maturity Index

As previously identified, there is an existing need of solutions such as the acatech Industry 4.0 Maturity Index. Companies are looking forward to progress towards Industry 4.0 and not only the framework that acatech provides, but the detailed roadmap, has caught the attention of several companies that have already decided to use it.\(^{19}\)

Being able to use the model with companies have allowed acatech to validate it, and also to analyse in which maturity stage were the participants. After this, common areas requiring action were identified.

3.2.1. Maturity analysis

Using the acatech Industrie 4.0 Maturity Index allowed companies to determine their current maturity stage. In Figure 19, there is an overview of all the stages of the journey, considering that Computerization and Connectivity are not part of Industry 4.0, but from the previous Industrial Revolution.

![Figure 19. Industrie 4.0 maturity stages. acatech.](image)

The vast majority of companies that participated in the study, resulted to be in the third maturity stage (Visibility). This means that they just started their journey towards Industry 4.0 and they are working on achieving a full visibility across the organization. Only a really small percentage

of companies had already achieved full visibility and was working on the next stage which involves understanding this visible data.

This fact is surprising because the companies that wanted to take part in this study where somehow involved with the Fourth Industrial Revolution and could be considered pioneers. But the reality is that, according to acatech, they have just started their journey.

### 3.2.2. Areas requiring action

The immature stage in which companies were placed, have made think about which are the main challenges that were avoiding them to make the next step. This lead to the identification of areas requiring action to unlock.

The most important topic requiring action for each functional area is listed below.

- **Resources.** Different and incompatible data protocols within resources that should be connected was found, not allowing to ensure communication standards.
- **Information systems.**
- **Organisational Structure.** Companies still have traditional organizational structures that do not allow fluid interdepartmental collaboration.
- **Culture.** There was not enough active employment involvement.
4. Analysis. Usage of Process Mining within Industry 4.0

The journey towards Industry 4.0 have been defined, and companies are starting to position themselves. In the current scenario, the digitalisation era have been surpassed, comprehending computerisation and connectivity stages.

In this situations, companies are now able to proceed looking forward maturity in Industry 4.0 by allowing the collection of absolutely any kind of data from the organization, and then by being able to understand this data. This comprises visibility and transparency maturity stages.

It is important to keep in mind that the Fourth Industrial Revolution does not only involve a technological development, rather a full organizational evolution encompassing, among others, a cultural approach.

This means that the outcomes of using Process Mining should not only be technological, but also help in the development of other structural areas.

4.1. Industry 4.0 Maturity Stage limitations

Process Mining is a relatively young technique that, since it has been born in a really technological advanced era, it requires of certain developments to become meaningful.

As described before, Process Mining works with event logs, which are data from the activities within a company. Whereas it is true that individual processes might be analysed, this use does would not be contributing to the Industry 4.0 development.

In this way, the company needs to be mature. It is not only that the company needs to have visibility regarding data. In addition, employees must be able to understand the situation, systems must be prepared to implement any change needed and organizational structure of the company should accordingly facilitate it.

4.1.1. Digital Twin

The Digital Twin (or Digital Shadow as previously defined) consists of a virtual replica of a whole organization.

Digital Twins enable an intelligent environment combining data from connected devices, business applications, and really any other data source to model, monitor and manage real-world environments.
The capabilities of Digital Twins can be easily transposed or applied to any industry, process, or environment and can ultimately keep a business operating smoothly.

When Process Mining is enabled to work with a Digital Twin, its possibilities are endless. The tool gets control over absolutely any activity within the company, since they all create a digital footprint that is being tracked.

With process mining, a process is virtually reconstructed using data from the event logs. These event logs represent the digital footprint of any action that happened within the company’s operations. In fact, this reconstruction becomes a digital twin of the real world process.

Any Process Mining technique, such as process discovery, conformance checking or process enhancement will have a maximum impact in the organization, thanks to the availability of all the data.

4.1.2. Cross-organizational scenarios

Usually, Process Mining techniques are applied within a single organization. Due to globalization and the presence of large company groups, there is the possibility to perform cross-organizational analysis.

In this case, it becomes interesting to check the current maturity stage for each of the companies involved in the analysis, because there is the possibility that one of them limit the results. In addition, it could happen that only the more mature companies benefit from this technology, since the less mature ones could not be ready yet to understand and apply the insights.

4.2. Available solutions

Since the emergence of Process Mining as an academic discipline, some researchers decided to start offering these solutions professionally to companies and third parties. In this way most companies that offer Process Mining solutions have been born.

However, the number of solutions available is really limited, with only a couple of dozen vendors. It should be added that large technology corporations have recently entered the market through the acquisition of one of these companies.

Generally, all vendors provide similar solutions for most use cases. The available services focus primarily on solving processes with similar features and functionalities.
Conclusions and Outlook

Companies are already aware about the benefits of reaching full maturity within Industry 4.0 and are on their way. The idealized version that stands at the end of this revolution is clear: an hyperconnected company automatically adaptable. The problem results to be that the steps required to achieve it have not been explored in detail yet.

The acatech Industrie 4.0 Maturity index have set an important milestone by providing this requested walkthrough. It has been already used by several companies and it has allowed to identify which gaps need more development. Certainly, most of the companies studied resulted to be in early maturity stages.

The study showed that digitalisation era has now become a standard, and companies are looking for the next steps that will allow them to become leaders in the market. In this context, Process Mining satisfy the requirements needed to fill this gap.

Despite of this, the current scenario is not as ideal as it was expected for the implementation of Process Mining techniques. Its maximum potential will be unchained once companies have a complete Digital Twin. It has been proven that in this case, Process Mining usage will significate to jump one whole stage in the maturity index.

In the case of companies that undergo the Industry 4.0 development from an early maturity stage perspective, its benefits are not that clear. The detailed applications of Process Mining techniques have not been covered in this project, hence it requires further research.
References

acatech, Plattform Industrie 4.0 (2019). Key themes of Industrie 4.0. Research Council of the Plattform Industrie 4.0


