Digital Manufacturing and Design Training Network (DiManD)

Grant agreement No 814078– H2020-MSCA-ITN European Training Network Grant

Deliverable 4.1

Definition of Industrial Barriers and Challenges to Context-aware Autonomous Systems with Respect to Current Practices

Requirements Specification for Distributed Manufacturing Systems

March 2021

Lead parties for Deliverable : UNOTT
Deliverable due date : M19
Actual submission date : M22
Dissemination level : Public

All rights reserved

Table of Content

1 Introduction and Context 4

1.1 Objectives of this Deliverable 6

2 A Definition and Requirements Model of Automation for Distributed Manufacturing 8

2.1 Autonomy Introduction 8

2.2 Objectives of Autonomy 10

2.2.1 Automation or Autonomy? 11

2.2.2 Application of Autonomy in Engineering Domains 12

2.3 Models of and Requirements for Autonomy 23

2.3.1 Key defining aspects 23

2.3.2 Existing Models and Level Definitions of Autonomy 29

2.4 Defining Model of Autonomous Systems Requirements and Features 33

2.4.1 Data, information and knowledge 37

2.4.2 Process 39

2.4.3 Interaction 44

2.4.4 Infrastructure 48

2.4.5 Self-x capabilities 50

2.4.6 Measurement performance 52

2.5 Model Discussion & Conclusions 56

3 A Review of Barriers and Challenges to Autonomy for Distributed Intelligent Manufacturing Systems 58

3.1 Introduction 58

3.2 Technical barriers and challenges 60

3.2.1 Data, information and knowledge 61

3.2.2 Process 63

3.2.3 Interaction 66

3.2.4 Infrastructure 74

3.2.5 Self-X 77

3.2.6 Measurement performance 79

3.3 Social, Economic, and Environmental Barriers and Challenges 81

3.3.1 Social aspects 81

3.3.2 Economic aspects 91
3.3.3 Environmental aspects 93
3.4 Barriers and Challenges Discussion & Conclusions 95
Bibliography 96

4 Versions **115**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

Chapter 1

Introduction and Context

Manufacturing today has to cope with increasingly dynamic and changing environments. Market fluctuations, the effects caused by unpredictable material shortages and supply chain disruption, highly variable market demands in terms of volumes and product types, and the availability of workers and availability of skills can all require system robustness and resilience.

A proposed solution to this type of variability is the employment of distributed intelligence in manufacturing systems - also called decentralised control. Traditional control strategies in manufacturing are highly centralised and hierarchical to optimise for fixed, high volume production. This however also leads to rigidity and an inability to rapidly respond to changes [1]. A variety of approaches have been proposed to improve the responsiveness of manufacturing, perhaps the most well known being the principles of lean manufacturing [2]. This involves a shift to a network model of organisation (instead of hierarchical) and enables workers to make local decisions and rely on dynamic self-coordination.

Lean manufacturing and related disciplines focus on human decision making in distributed complex systems. However, technological development in machine intelligence and communication standards is starting to replicate the dynamics of empowered workers making distributed local decisions in the digital domain. Digital intelligent manufacturing - commonly called Industry 4.0 (or I4.0) [3] - is a drive to digitalise manufacturing processes and enable vertical and horizontal data sharing. This shared ubiquitous data can then be used to coordinate manufacturing operations, be used for monitoring and optimisation of processes, and enable the use of artificial intelligence methods to discover new insights. As a concept, I4.0 ties together a lot of modern thinking about how manufacturing will look like in the immediate future.

Two key related enablers of Distributed Manufacturing and Industry 4.0 are the Internet of Things (IoT) and Cyber-Physical Systems (CPS). IoT (sometimes called the Industrial Internet of Things [4] or IIoT in the manufacturing context) is integrating computational and networking capability into physical "things". In the manufacturing context this means enabling the manufacturing assets (and even the products - see Deliverable 4.2) to be intelligent, networked, and collaborative. CPS (sometimes called Cyber-Physical Production Systems or CPPS in the manufacturing context) is a physical manufacturing system where the behaviour is controlled by software intelligence [5]. Unlike traditional embedded systems, the behaviour emerges from a collaborative network of devices rather than each device being stand alone. A true CPS has such closely integrated physical and software components that the two cannot be considered independently. The use of these two approaches enable distributed manufacturing intelligence at all levels of a manufacturing system, including 'at the edge'. This enables localised decision making and data processing, as well as data exchange and networking to allow assets to share data and collaborate.

These new developments in embedded, networked computing in manufacturing offer multiple opportunities to meet the dynamic demands and changing environments faced by modern manufacturing enterprises, by enabling rapid localised analysis of data and response to disruptions without unnecessary escalation. Distributed intelligence and decentralised control in manufacturing are not new concepts. The use of decentralised manufacturing control (e.g. via the use of multiple Programmable Logic Controllers) has been explored [6], particularly in the context of re-configurable manufacturing systems [7]. However, a key modern challenge for Distributed Manufacturing is the concept of *Autonomy* - the ability for a system to make independent decisions and respond to new unexpected situations without external input. For a manufacturing system with distributed intelligence to truly meet the demands of dynamic manufacturing environments it must respond to new and unforeseen situations - and hence must be autonomous.

Autonomy is being adopted in different industrial contexts and research domains - it is a key enabler in the field of self-driving cars for example [8]. However, there are clear divergences when describing the concept of autonomous systems and what autonomy actually means as a term. To realize the implementation of autonomous behaviour in manufacturing systems it is essential to specify what is exactly meant by autonomy, how autonomous manufacturing systems are different from traditional and other manufacturing systems, and how autonomous manufacturing systems can be identified and achieved based on their features and enabling technologies. It is also required that the barriers and challenges preventing companies from implementing these technologies and approaches are identified.

Contextual note: "Distributed Manufacturing" has two possible definitions within manufacturing research. The first relates to the distribution of intelligence in a manufacturing system and is largely synonymous with the decentralisation of control i.e. the transition from a single central control system to delegated control distributed amongst the assets which comprise the manufacturing system. The second definition relates to the business and supply strategy of geographical distribution of manufacturing facilities, using multiple smaller, leaner production units rather than single monolithic factories. To exacerbate the issue, both areas of research contribute to greater flexibility and responsiveness of manufacturing in the face of disruptions.

This deliverable, task, work package, and project use the *first* definition of distributed manufacturing, as befits the project focus on digital manufacturing. The term ”Distributed Manufacturing Intelligence” is used to help disambiguate the terms.

1.1 Objectives of this Deliverable

This document represents deliverable D4.1 - *Requirements Specification for Distributed Manufacturing Systems.* It is the only deliverable of Task 4.1 - *Definition of Industrial Barriers and Challenges to Context-Aware Autonomous Systems with Respect to Current Practises*, which is part of Work Package 4 - *Autonomous Context Aware Manufacturing Platforms.* The objective of the work package is to enable increased adoption of distributed, intelligent, autonomous manufacturing systems which include the products themselves as active assets within the production process.

Distributed intelligent manufacturing requires three key elements: the enabling architecture on which system intelligence can operate (i.e. the IoT and CPS considerations), an understanding of what intelligence a system needs to exhibit in order to maximise productivity and other KPIs (i.e. the distributed autonomous intelligence), and the technical implementation by which the distributed intelligence can be achieved (e.g. a multi-agent system).

The requirements and challenges regarding the enabling technical architecture for distributed manufacturing systems will be developed in Work Package 3. The requirements and challenges regarding the agent-based implementation of the distributed intelligence will be developed in Work Package 5. To ensure work is not unnecessarily replicated, this task in Work Package 4 focuses on defining what the requirements for distributed autonomous intelligence is, and the barriers preventing its adoption.

This deliverable divides this challenge into two parts:

1. Definition of what ”Autonomy” is in this context and the requirements for its implementation. To progress the field of distributed manufacturing, clear requirements for how autonomy could be implemented must be defined. Firstly however, there is a lack of a clear definition of autonomy, and a lack of a clear model for how autonomy could be approached and implemented. We introduce a definition of autonomy for the manufacturing context, and a model that represents the incremental requirements for introducing a distributed manufacturing system and then progressing to full autonomy.

2. Identification of the barriers and challenges which are inhibiting the uptake of autonomous systems in manufacturing, both in terms of technical barriers but also the social, economic, and environmental barriers.

The remainder of this document comprises two chapters - one each for the two points above.

Team Involved in Deliverable Writing

ESR1: Fan Mo, The University of Nottingham
ESR2: Agajan Torayev, The University of Nottingham
ESR4: Ngoc Hien Nguyen, Mondragon Unibertsitatea
ESR7: Jose A. Mulet, STIIMA-CNR
ESR8: Fabio Marco Monetti, KTH Royal Institute of Technology
ESR9: Sylvia Nathaly Rea Minango,KTH Royal Institute of Technology
ESR14: Hamood Ur Rehman, The University of Nottingham and TQC Ltd
Supervisor: Dr Jack C Chaplin, The University of Nottingham

Chapter 2

A Definition and Requirements Model of Automation for Distributed Manufacturing

2.1 Autonomy Introduction

Manufacturing today requires an increasingly dynamic approach with rapid, diverse changes to meet the demands of the market. Short product life cycles, as well as decreasing batch sizes with simultaneously increasing product variants and complexity, are challenging the traditional production systems [9, 10]. Current conventional structures and methods cannot handle changes, unpredictable events, and disturbances in a productive, cost-effective manner [11]. To manage these dynamics, the new industrial concept of Industry 4.0 has emerged, a trend linked to digitalization and distributed intelligence in systems which could enable factories to achieve higher production variety with reduced downtimes [12, 13]. These aspects are enablers for distributed industrial systems which can handle increasing complexity [10, 14], and respond quickly to unexpected disturbances without the need for centralized control and re-planning and with minimum human intervention [15, 16, 17], while improving yield, quality, safety, and decreasing cost and energy consumption [14, 17, 18]. The ability to respond to these unplanned distrubances and events is typically called *Autonomy*.

For these systems to be successfully introduced into the manufacturing industry, it is essential to establish what autonomy means in this context. As the concept of autonomy has been associated with independence [19] and living entities [20], the first definitions of autonomy come from outside the engineering domain. Kant's philosophy defines an autonomous being as a self-governing individual [21]. This notion of autonomy and its relation to biological individuals are found in autonomous organisms that not only reacts to stimuli, but makes decisions based on its internal constraints, builds its own identity, and transforms the environment as a result of its actions [20].

As the concept of autonomy has drawn attention from different research fields, several definitions have been used within different domains. However, this increasing adoption of the word “autonomy” to define or qualify particular types of systems has not provided a single unified structure of characteristics that can be used universally to distinguish a system as “autonomous” or not. The research literature shows clear divergences when describing autonomous systems. Specifically, in the context of manufacturing assembly systems, B. Scholz-Reiter et al [11] mentioned autonomy as the independence of a system in making decisions by itself without external instructions and performing actions by itself without external forces. However, this definition is often confused with the concept of automation systems that are designed to carry out independently tasks and even process local information along pre-programmed supervised tasks without human intervention [14][22]. In an effort to distinguish autonomy from automation in the chemical process industry, Thomas Gamer et al [14] indicated that the act of progressively handing over more and more of these automation processes to the system is defined as a gradual transformation to autonomy, which does not take into account the self-governing characteristic of autonomy according to Kant’s philosophy. In the field of logistics, Katja Windt et al [9] pointed out that the ‘control’ aspect of autonomy in logistics systems is commonly used to refer to the ability of logistic objects to process information, and to render and execute decisions on their own. In the context of a machining manufacturing system, Hong-Seok Park et al [15] defined the broader ability of autonomy as an ability of human, robots, or software to achieve its goals without any support from the others while the interaction among them represents social ability of the system to achieve its global goal.

There is a lack of universal consistency regarding the definition of autonomy, its features and requirements, and how they are qualified in different research and engineering domains. High-level text definitions of autonomy may exist, but they lack the precision required and do not break autonomy down to define the requirements. Some models have been made to qualify and capture autonomous completeness based on autonomous ’levels’ which are defined as the extent at which autonomy exists on a continuum ranging from no autonomy to full autonomy [23, 24]. These models come closer to a functionally useful definition of autonomy for distributed manufacturing intelligence. One of the models proposed by Beer et al. [23] uses five stages on the assessment of the whole factory, but a more detailed description of the levels of autonomy is required to support the industrial application and transition [25]. This is especially true in the manufacturing context as the implementation of autonomous systems could have a substantial impact on productivity and cost reduction [10, 26].

Therefore, this work aims to offer a unified definition of autonomy in the manufacturing context as a framework for both industry and research. This is achieved by: (1) providing an analysis on the state of the art regarding autonomy definitions in different manufacturing contexts (2) identifying similarities and differences in the definitions, (3) defining the common features of autonomy, (4) and harmonizing these features into a five-stage model of autonomous manufacturing. This model defines autonomy from the manufacturing perspective and comprises enabling features associated with maturity levels providing a comprehensive appraisal of manufacturing systems from the as-is stage, through distributed manufacturing approaches, and to full autonomy.

The following section (2.2) will clarify the concepts of automation and autonomy and then discuss how autonomy is portrayed in different engineering perspectives. Section 2.3 presents and analyzes the key requirements and features of autonomy, which subsequently forms a five-stage model defining levels of autonomy associated with maturity levels for each autonomous feature in Section 2.4. This model is the definition of autonomy for manufacturing, and aims to facilitate the complete identification and evaluation of autonomous manufacturing systems. Finally, the last section (2.5) presents the conclusions and future work perspectives with regards to the model.

2.2 Objectives of Autonomy

In general terms ‘Autonomy’ can be defined as the capability of individual entities to collaborate, implement, and direct towards achievement of a specific goal without any external influence [9] under its own laws and control principles [27]. In etymological perspectives, this concept is embedded in the aptitude of individuals to take independent decisions or self-government (*i.e.* the ability of one to dictate oneself his/her own law) [27].A conjecture in this view is that higher the level of uncertainty and the ability of the system to adapt to it, the higher the level of autonomy.

[28] defines autonomy as a design goal that can be considered as the ability of a system to function independently, subject to its own laws and control principles. [10] give the definition of autonomy as the capability of an entity to create and control the execution of its own plans and/or strategies. According to [15] autonomy allows the system to respond and recover without modifying scheduling. Autonomy is an ability of the agent to achieve its goals without any support from the other agents. [16] describe autonomous systems as intelligent machines that execute high-level tasks without detailed programming and without human control. They know their capabilities and their state. They are able to decide between a set of alternative actions, orchestrate and execute skills. Autonomy is defined as the ability of an entity to structure its own action and environment independently and without unwanted influence from the outside [29]. It is the ability of a system to make its own decisions and to act on its own, and to do both without direct human intervention [30].

Some authors have developed a scale to try to measure the degree of autonomy quantitatively [29] by measuring the number of decisions taken by the system or entity on its own, or by establishing levels of autonomy according to different criteria accomplished by the system [9]. Based on those characteristics, a high-level definition of autonomy could be assembled for the manufacturing field as the capability of an entity to locally make decisions derived from reasoning on its current situation and previous knowledge, planning a strategy, and executing necessary measures to achieve the desired outcome.

2.2.1 Automation or Autonomy?

As seen above, many authors give a general definition of autonomy including a key element of the ability of the entity to achieve its goals without external influence, be it human or from other agents. However, this can and sometimes does create confusion between the concepts of automation and autonomy.

Automated devices execute actions determined by a predefined set of rules to produce an outcome, with minimum or no human intervention [26, 31], so only foreseen outcomes resulting from the execution of a plan based on sequential logic are possible [17, 19, 26, 31, 32]. It is expected that if these devices deploy in an unpredictable environment, their fixed plans will rapidly encounter limitations. Autonomous individuals (commonly human workers) which exhibit higher intelligence and flexible behavior would be required in these situations, that is the reason why autonomy is understood as one step beyond automation [19, 33, 32].

The key differentiator is the ability for an autonomous entity to face unanticipated situations and adapt its course of action as they appear [20, 31], which is enabled by their self-determination and cognitive capabilities [34]. Domain knowledge acquired through learning provides them with implicit intelligence [17, 26, 33], and a high level of understanding [31, 26] and allows them to take conscious decisions to accomplish their goals [33]. Thus, the closer an entity approaches full autonomy, the smaller the role of any external agent in its operation will be (Figure 2.1). In the specific case of humans, this could become a two-edged sword because as the system becomes more independent and reliable, the less aware and prepared the operator would be to take over control, which is referred by [17] as a critical barrier to autonomy.

Latest technological developments (e.g. autonomous driving, unmanned aerial vehicles, and artificial intelligence) have raised critical concerns about the extent to which autonomy should be developed or even allowed in systems, and the same concern applies to the participation of autonomous entities in production systems [31, 33, 32]. Defining the upper limits of autonomy in intelligent systems is a complex task at the current technological stage, as the challenge is to develop an autonomous system in which the human would still be able to regain control in exceptional situations [35], that does not represent a hazard, and which "can co-exist with people, and be trusted, safe companions and co-workers" [26, pp.3]. Additionally, the intelligence and self-government characteristics that a fully autonomous system could achieve raises concerns beyond human safety to more complex ones like liability, moral accountability, ethics, decision responsibility, privacy, and security issues [36, 37].

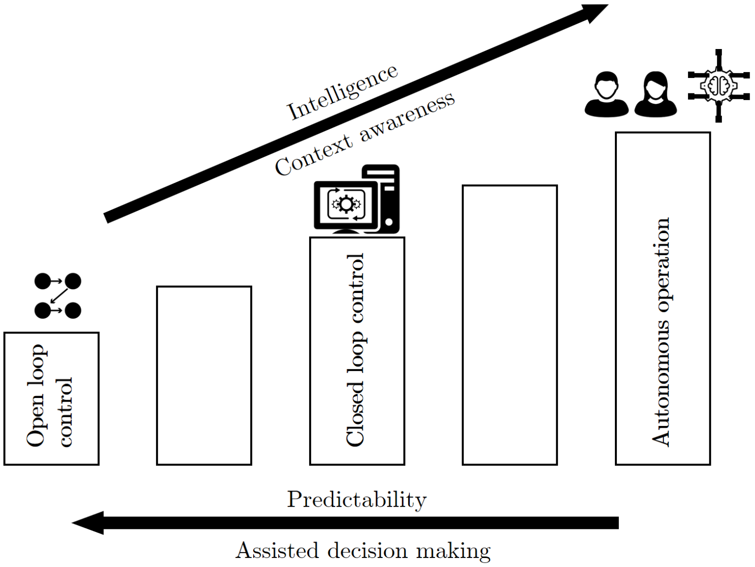


Figure 2.1: Progression from automation to autonomy

2.2.2 Application of Autonomy in Engineering Domains

The term autonomy has also been quite fluidly used in industrial engineering applications beyond manufacturing, and is generally used to describe a means of controlling and directing a certain functionality of the system without human intervention. A significant rise in autonomy applications in engineering applications have been witnessed, and the development of autonomous vehicles in particular, autonomous robots, maintenance, processes and quality systems has seen major progress in the research. We discuss this developments here to determine what progress can be applied to the manufacturing domain.

Common motivations for autonomous industrial applications lies with economic, performance, and human safety aspects [33]. [33] argues that such possibility of autonomy in automation is only followed with a certain degree of 'consciousness' i.e. ability of the system to be 'aware' of its environment and context along with a deliberation capability of its information processing priorities.

Unmanned Aerial/Underwater Vehicles

The application of autonomy has been especially significant in the area of unmanned vehicles [38]. The very first Autonomous Underwater Vehicle (AUV) came into existence through the work of Washington University [39]. The Self-Propelled Underwater Research Vehicle (SPURV) was designed for underwater exploration and acoustic communications. Functionality has been expanded with higher sensor payload capacities and longer endurance while controlling the container size. A Norwegian team has developed an AUV for detecting gas leaks and chemical underwater discharges while keeping the solution cost-effective [40]. The application of autonomy in underwater vehicles has expanded to the areas of coastal area monitoring, scientific explorations, military applications, offshore mining and seafloor mapping [41][42][43] .

Beyond underwater applications there have been significant strides in autonomous aerial vehicles with examples helping to investigate changes in Antarctic ice shelfs which has a direct relation with climate change and global warming [44]. In general aspects the autonomous flying functionality in vehicles is actually remote controlled drones with pilot assistance features. Boosting flight performance, protection in hazardous and risky works, and cost optimization [45] have been popular areas of autonomy application in vehicles. In all of these applications, barriers and challenges include limited applied autonomy, and operational bandwidth [46] which limits communication to a specific range or at the start and end of operation. Expensive maintenance and programming for highly reactive decision making in uncertain environments [47] are prominent drawbacks in autonomy realization in this areas. Time and energy restrictions are other problematic aspects that the machines should wisely manage by controlling their resources to achieve extended autonomy [47]. In the case of unmanned aerial and underwater vehicles the concept of autonomy is primarily the main driving factor for further research [38].

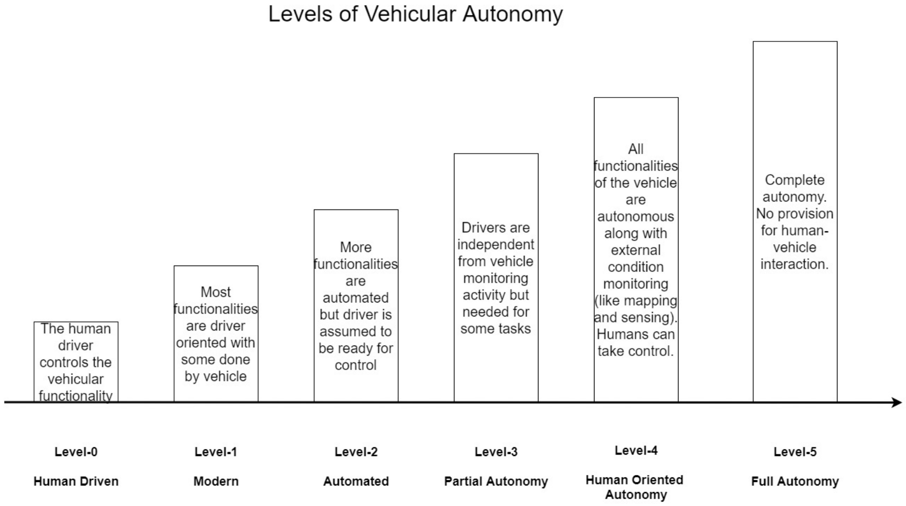

Levels of Vehicular Autonomy
The human driver controls the vehicular functionality
Most functionalities are driver oriented with some done by vehicle
More functionalities are automated but driver is assumed to be ready for control
Drivers are independent from vehicle monitoring activity but needed for some tasks
All functionalities of the vehicle are autonomous along with external condition monitoring (like mapping and sensing). Humans can take control.
Complete autonomy. No provision for human-vehicle interaction.
Level-0
Human Driven
Level-1
Modern
Level-2
Automated
Level-3
Partial Autonomy
Level-4
Human Oriented Autonomy
Level-5
Full Autonomy

Figure 2.2: Transition from Human Driven to Self-Driven- A simplified representation of SAE J3016 Automation Levels [36]

There has been significant rise in demand of autonomous vehicles that can be smart enough to not only manage their own functions but also interact with other equipment, monitor and store information while maintaining sustainable working conditions [36]. Initially the role of autonomy in these use cases was focused on removing the need for a local operator thereby subjecting these equipment to much harsher environments than possible by humans [45]. There exists significant gap between ensuring full autonomy in vehicular systems and the control system required to implement it [46]. A limitation of research done in this area is maintaining the focus on one or more aspect/feature of autonomy while ignoring others.

Figure 2.2 shows the SAE J3016 Automation Levels which define the levels of autonomy in vehicles, from human operated to self-driven. The levels are defined from 0-5 with 'Level-0' being associated with human-driven to 'Level-5' where no interaction happens between human and vehicular control. The most important aspect in the figure is the transition from Level-3 to Level-4 where the balance of autonomy moves from humans to the vehicle itself

Autonomous Robots

Most robots such as industrial robots (KUKA, ABB, Fanuc, among others), humanoid robots (e.g. Atlas[48], Archie, Asimo), consumer robots (such as automated vaccum cleaners), disaster response robots, and many others require energy management as these are typically battery powered. [49] established a case of energy autonomy in autonomous robots developed from the concept of 'potential energy', where actions are constrained by remaining energy and relational distance among individual robots. The robot as an adaptive 'agent' (i.e. an autonomous entity) reacts to the changes in the environment while cooperating in a network of robots to achieve energy autonomy.

The energy conservation strategies in autonomous robots presents a case for the employment of multi-agent systems [50]. Each agent in this system has a belief-desire-intention (BDI) software architecture where each agent is responsible for energy conservation as a 'desire' and co-ordinates with other agents to achieve this goal. It could be argued that if desires are predefined in the system then does this make the system autonomous? An autonomous system in the case of robots could develop its own desires and should have the authority to enforce it by evaluating the environment it interacts with [51]. Realistically an autonomous robot can only be truly autonomous if it achieves an outcome with minimal external assistance.

The goals that drive the robot or control its behavior are associated directly with the main constituents of a robotics system - perception, actuation, and decision making (Figure 2.3)[52]. 'Perception' is concerned with observing the environment with equipment like laser scanners, vision cameras, and other sensors that serve to input the current environment state into the system. 'Decision' involves locally selecting the best operation to match the goal of the system. 'Actuation' involves equipment such as robotic arms, wheels or tracks, or other manipulators that allow the robot to act on the decision and change its environment to achieve its goals.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

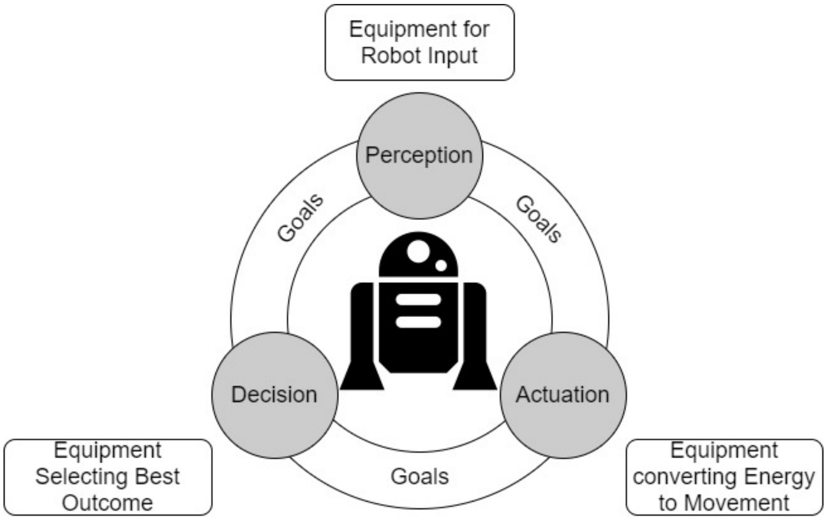

Equipment for Robot Input
Perception
Goals
Goals
Decision
Actuation
Goals
Equipment Selecting Best Outcome
Equipment converting Energy to Movement

Figure 2.3: The main constituents of autonomous robots - perception, decision,and actuation all driven by the robot's goals

Autonomous Maintenance

Autonomous maintenance reduces the need for human intervention in maintenance operations for dangerous and dirty environments, and helps improve the perception of the task to one of cost savings and productivity [53] . Automated maintenance is different to automation within manufacturing as the maintenance operations possess a different set of characteristics. Maintenance operations are considered to be non-deterministic (age, usage or other factors of the asset can affect the duration between required maintenance), usually non-uniform (multiple failure modes and the condition of the asset will define the actions required) and highly irregular (varying levels of unplanned and planned activities) which makes the scheduling of maintenance for autonomous equipment especially difficult.

The application of autonomy in maintenance operations is geared towards non-destructive testing (NDT) and inspection (NDI). Pipe Inspection Gadgets (PIGS) offer some level of autonomy during industrial maintenance of pipes, as pipelines can restrict wireless communication [54]. Some solutions have been developed to detect cracks along pipe infrastructure that utilizes fully autonomous robots capable of performing assigned tasks autonomously [55][56]. Aircraft oil delivery tube crack detection has also seen employment of autonomous robots for routine inspection of fatigue cracking [57] by using magnetic flux leakage and eddy current array techniques. Movable pipe robots have been developed that utilise electromagnetic stabilization to grip surface before inspection [58].

Hazardous environments such as nuclear plants require mobile autonomous equipment [59][60].Such robotic applications have also been extended to ship-hull welding, steel bridge inspection, and nuclear power plants fault detection[61][62][63]. Building maintenance and surveillance is an important but mundane task that is a clear opportunity for autonomous systems, and autonomous robotic maintenance systems have been developed wherein robots roam throughout the building, use RFID tags to collect and relay information to centralized maintenance center thereby increasing efficiency and reducing labor costs [64]. Such is also the case with power distribution line maintenance where autonomous robots were developed to perform the task [52]. The robot receives an instruction from the operator and plans its trajectory to perform the task using image processing, feature recognition, and planning algorithms.

The availability of historic data to make informed inferences is the primary enabler of autonomous maintenance [65]. Correct data is essential for drawing accurate results. The historical insights gives the likelihood a of a state, and a possible distribution for future predictions.A huge research area explored in this domain is to converge the gap between the estimated, predicted, and actual states of the equipment [66]. This will help in predictive and remaining useful life (RUL) estimations.

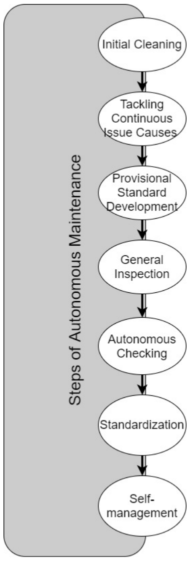

Steps of Autonomous Maintenance
Initial Cleaning
Tackling Continuous Issue Causes
Provisional Standard Development
General Inspection
Autonomous Checking
Standardization
Self-management

Figure 2.4: Steps of Autonomous Maintenance with Total Preventative Maintenance

Total Preventive Maintenance (TPM) [53] outlines the main steps to complete maintenance operation. ‘Initial cleaning’ involves dealing with equipment problems and performing actions to eradicate current existing equipment issues (like lubrication and tightening). ‘Tackling Issues’ goes upwards a level by focusing on the issues that contribute to the problems and reducing them by implementing formal processes. Based on these issues tackled the next step is driven by ‘provisional standard’ development for in-house maintenance at periodic intervals followed by empowering the automated equipment to inspect and monitor itself. This introduces the concept of autonomy in maintenance application. The compliance with international standards for maintenance enables accurate and compatible data recording, material movement, control and quality assurance. This in turn leads to 'self-management' (Figure 2.4).Self-management in maintenance are derived as actions that could be controlled by machine/equipment itself to improve its condition, maintain or restore to a specific condition.

Autonomous Processes

Research in autonomy in the process industry has taken inspiration from research in the self-driving automotive domain to develop a definition of autonomy focused on plant operators [67]. This is further refined by work of [68] into a taxonomy of autonomous key features. Smart manufacturing has been seen as an enabler for digitalizing and automation in industries like chemical, petrochemical, and pharmaceuticals but the lack of modern cyber-physical systems, interoperability, and flexible architectures present challenges to its application [69][70]. [71] builds on this foundation and details the challenges in the smart process industry as well as into topics of communication, interconnection, uncertainty, robustness and, predictability from data. The works present the case that opportunities and challenges are not limited to production only but to the extended supply chain.

Autonomy in this application is used to maintain the effectiveness of continuous process [14] despite variation and uncertainty in the supply chain [71]. An autonomous process may fail to act to its full capacity if it gets restricted due to a lack of materials or resources, and research focuses on improving this resilience and robustness.

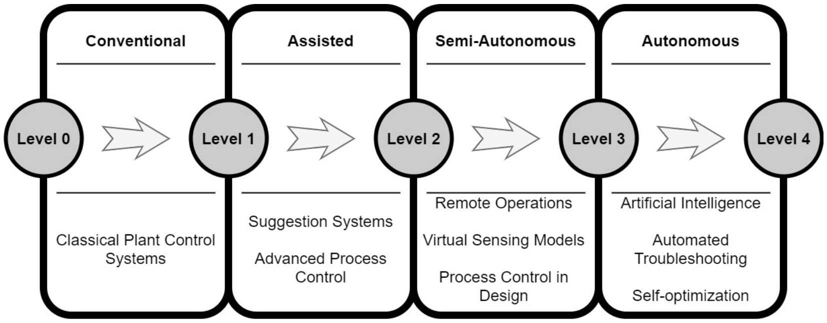


Figure 2.5: Levels of autonomy in process industry from conventional to autonomous process

Figure 2.5 defines levels of automation in process industry from Level-0 to Level-4. Level 0 has no automation presence. As the automation level increases the control of process parameters increases and becomes more intelligent, and is enhanced by advanced process control and expert systems. The semi-autonomous systems permit remote operation and use of digital twins.

Full autonomy is achieved at Level-4 by integrating artificially intelligent autonomous decision making and control for self-optimization, troubleshooting, and self-management of the process plant.

Autonomous Quality

Intelligent quality systems derived from work of [72] utilize networked machines capable of formulating quality functions and automated analysis for improving quality in production systems. ‘Quality 4.0’ builds on the increasing digitalisation of manufacturing and applies the Total Quality Management (TQM) approach enhanced with artificial intelligence, machine and deep learning, and becomes integrated with statistical process control [73]. [74] describes how modern quality systems can utilize seven tools and technologies for quality management. These are data science and statistics, enabling technologies (Industrial Internet of Things, integrated systems, cloud computing, and augmented/virtual reality), big data, blockchain, machine learning, neural networks, and deep learning.

The case of automated quality has gained significance especially with development of machine learning and deep learning models [75]. Machine vision integrated with quality control has also brought about significant impact to traditional go / no go scenarios [76]. The modern quality systems are increasingly utilising autonomous infrastructures to maintain strict statistical process and quality control (SPQC). Quality assurance frameworks that enforce quality control standards are required to be built-up to align themselves with strict codes of autonomous quality control [74]. Increasing competition in the market and demand for customized products has required manufacturers maintain increasingly strict quality standards aided by autonomous systems to produce high quality varied products.

Figure 2.6 shows how autonomous quality control influences different phases of manufacturing execution. The more advanced the level of autonomy, the more predictive quality control in the production process can be performed. Preventive maintenance is the regular upkeep and calibration of processes to prevent failures and maintain quality , reactive maintenance is the adjustment of processes as they are performed to maintain quality, and predictive is adjusting processes for the process begins to maintain quality. The earlier the quality control is performed, the fewer parts will need to be scrapped or fixed, improving productivity [74].

Autonomous Information and Communication Technology (ICT)

Deep learning has contributed to autonomous system development through rapid developments in computer vision, acoustics, tracking, natural language processing and pattern recognition [75]. Improved computing capabilities has enabled the increasing accessibility of advanced machine learning models and processing. A good representation is the rapid progress in the development of unmanned ground and aerial vehicles and medical application robots due to new powerful computation frameworks like Caffe, Theano and Tensorflow

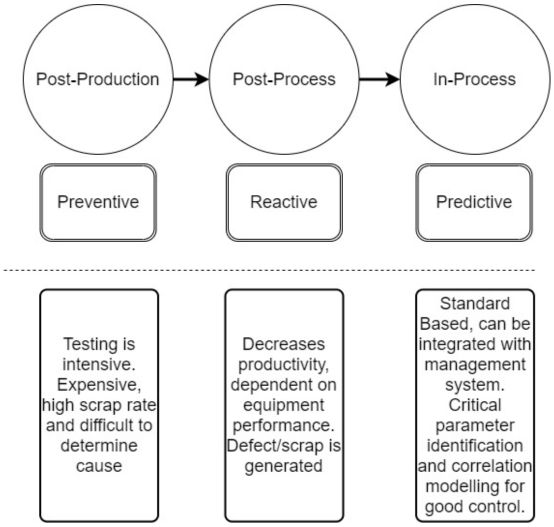

Post-Production
Post-Process
In-Process
Preventive
Reactive
Predictive
Testing is intensive. Expensive, high scrap rate and difficult to determine cause
Decreases productivity, dependent on equipment performance. Defect/scrap is generated
Standard Based, can be integrated with management system. Critical parameter identification and correlation modelling for good control.

Figure 2.6: Autonomous Quality Control at different manufacturing phases: preventive, reactive, and predictive

[77]. These have helped in building novel and robust unmanned autonomous systems. Machine learning capabilities provide perception and control capabilities that can mimic human interaction. These learning domains are aided by sensory perception guided by models to transform information to abstraction levels that map the environment [78]. Object detection [79][76], classification [80] and semantic understanding [81] relate to the vision abstraction that utilizes convoluted pooling of neural network layers for a multitude of machine learning tasks. These methods generally promote the independence of entities within its system and associated systems [82].

The development of new learning algorithms and applications have brought about a push in the digitalisation of industrial sector, particularly for areas and problems which were previously considered too complex to automate. This push in digitalisation of industrial sector requires the development of common standards and management frameworks and architectures however. ICT-enabled manufacturing management frameworks like ARC Advisory Group's Model [83] and the International Society of Automation's ISA95 Enterprise-Control System Integration hierarchical models [84] have been developed to meet these requirements [85]. These Architectures, standards and protocols for autonomous systems lay the foundation for autonomy implementation [86], but a pivotal gap in this domain is the lack of practical proven applications of these developed paradigms.

Distributed Manufacturing

On the production floor, distributed manufacturing intelligence can reduce setup times by planning or through reactive control, and reduce machine breakdowns by implementing TPM with machine data acquisition and status information monitoring. [18]. Proper tracking and tracing of objects (with autonomous identification technologies, proper marking, and storage and buffer size allocation), effective warehouse management (enabled by Manufacturing Execution Systems (MES) and Enterprise Resource Planning (ERP) systems), and fast fault detection (a combination of software checking of data values and statistical process control) are further aspects of this approach that benefit the manufacturing shop floor [18].

The rise of modular, flexible, and reconfigurable manufacturing systems has led to further developments in distributed intelligence such as holonic manufacturing environments [87] and the autonomous adaption of manufacturing systems to unexpected changes [88]. Multiple architectures such as ADACOR (Adaptive Holonic Control Architecture), PROSA (Product Resource Order Staff Architecture), HCBA (Holonic Control Based Architecture), ORCA(Orchestration and Reconfiguration Control Architecture) and POLLUX. have been developed that target improved flexibility, emergent behavior, new application domains, connectivity, and adaptability [13].

To realise autonomy within a distributed production environment the facility should support dynamic operations, dynamic prioritization, task migration and hierarchical restructuring [89]. Manufacturing system monitoring is essential to ensure quality requirements are met, the minimization / maximization of objectives is achieved, and faults are diagnosed and fixed. An autonomous facility would require effective control principles with an adaptive capability for accuracy, fault detection and identification [28].

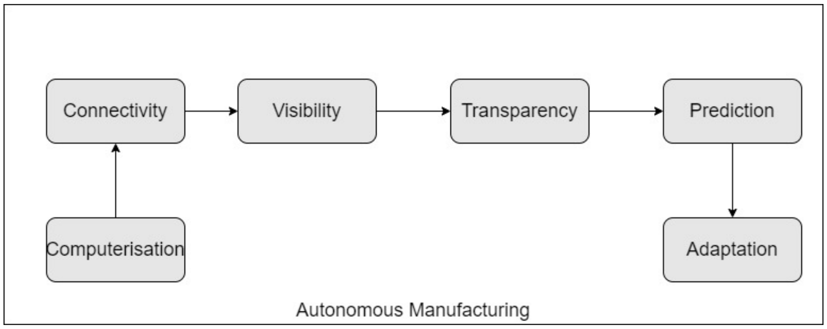


Figure 2.7: Steps to Autonomous Manufacturing from basic computerization of tasks to adaptation of digitisation in manufacturing industry

Production systems are assumed to be autonomous if they are able to perform their tasks without external influence [90]. However, a production system is a group of associated processes [69] and these processes collectively produce an output by having raw material or parts passed through each processing stage. Each of these processing stages may be a single production step with a single machine producing a finished product or may be a group of processes cumulatively involved in such an endeavor [18]. If each station simply performs a task independently of the others this is decentralised manufacturing control. If the stations share data on their status to make local decisions this is distributed manufacturing intelligence. However, true autonomy can only be achieved if autonomous control is established at a 'higher' level with a view of the whole system. Autonomy is not only restricted to processes but also extends to the components coming and leaving the system, and some of the components or processes may also be sub-contracted. To ensure complete synchronous resilient autonomy it is essential to maintain end to end autonomy [91, 15], and a major challenge in manufacturing is maintaining such level of autonomy all at levels. Production process are often highly distributed [87, 88]: some parts may be produced at one facility while other at completely different in another region (this is the second definition of distributed information - see contextual note in Section 1).

Case of Autonomy by Objective

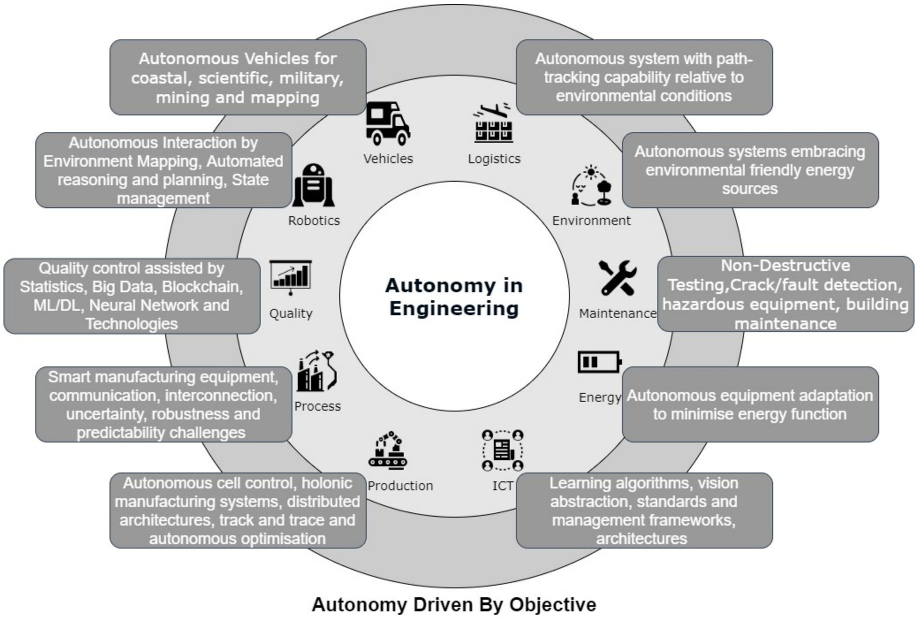

Autonomous Vehicles for coastal, scientific, military, mining and mapping
Autonomous system with path-tracking capability relative to environmental conditions
Autonomous Interaction by Environment Mapping, Automated reasoning and planning, State management
Autonomous systems embracing environmental friendly energy sources
Vehicles
Logistics
Robotics
Environment
Quality control assisted by Statistics, Big Data, Blockchain, ML/DL, Neural Network and Technologies
Quality
Autonomy in Engineering
Maintenance
Non-Destructive Testing,Crack/fault detection, hazardous equipment, building maintenance
Smart manufacturing equipment, communication, interconnection, uncertainty, robustness and predictability challenges
Process
Energy
Autonomous equipment adaptation to minimise energy function
Production
ICT
Autonomous cell control, holonic manufacturing systems, distributed architectures, track and trace and autonomous optimisation
Learning algorithms, vision abstraction, standards and management frameworks, architectures
Autonomy Driven By Objective

Figure 2.8: The business objectives driving autonomous behaviour in engineering applications; the figure gives an overview of research and trends in autonomous engineering applications

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

Table 2.1: Features and required functionality of an autonomous production system - Further explained in Section 4

Features	Description
Data Management	Structure, data processing, and integration capability of production systems
Knowledge Management	Ability to manage, deploy, and exploit acquired knowledge
Interoperability	Hardware and software compatibility
Synchronisation	Material and logistical physical synchronisation in system
Functionality	Mapping from Input (X) to Output (Y) and moving from explicit to implicit events for reaching autonomous behavior.
Optimization	Maximizing an objective function, typically increasing productive and reducing unproductive contribution
Decision Making	Moving from human decisions to independent decision making and execution
Reliability	Ability of system to maintain designed functionality and fault recovery
Context Awareness	Self-awareness of the system's physical, virtual and user environment
Interaction	Interaction between components, including parts, machines and humans in the production environment
Safety and Security	Protection of the system and humans from threats
Connectivity	Production system's ability to connect its elements to different network types (centralised or decentralised)
Control	Management of the operation of production system
Adaptability	The ability of a system to adapt to sudden changes

Configuration	Management of the system structure, including aspects of modularity, scalability and convertibility
Learning	Ability to evaluate information and experiences, and direct future behaviour
Accuracy	Identification of true positive/true negatives from false positive/false negatives
Stability	Ability of a system to maintain pre-defined standards and expectations in different environments
Certainty	Ability to be certain of measurements of operation

In all of these examples of autonomy research and practice, the final outcome (i.e. the objective) is the component that drives the use of intelligence in the system and the system can be compromised in the favour of result.If such a case the system has to be remodeled to map the desired outcome. This is mainly due to lack of standards and frameworks present to support the system thus maintaining integrity while achieving consistent results.

Different autonomous application will require different features to achieve their objectives. For autonomous vehicles the objective is the ‘functionality’ as the basis of autonomic behaviour. ‘Efficiency’ and ‘interactions’ are the most important aspects of autonomous robots. Autonomous maintenance on the other hand anchors the use of ‘historical data’ and ‘measurement’. Main driver for autonomy in process applications are ‘adaptability’ and ‘repeatably’ whereas in quality the autonomy is catered in the process to ensure ‘accuracy’ and ‘reliability’. ICT aspect deals with the ‘knowledge’ attribute of autonomy. Manufacturing further builds on the concept with focus primarily on end-to-end ‘synchronisation’ thereby controlling the manufacturing process. These derived features along with others lead to our general concept of ’Autonomy in Manufacturing’ (Table 2.1) discussed in Section 4.

2.3 Models of and Requirements for Autonomy

Autonomy is becoming more and more important. In this section, models and requirements of autonomy will be introduced

2.3.1 Key defining aspects

Autonomy is a multidisciplinary concept, and accordingly a multidisciplinary approach must be considered to achieve a successful deployment of autonomous system in a manufacturing context. Autonomy covers aspects from digital technological developments, to business concerns. The main key defining aspects and requirements to increase robustness against environmental complexity and dynamism are defined by[18] as: i) Humans and their interactions with the environment and systems; ii) Available machines and knowledge of the processes; and iii) Software entities and algorithms for decision-making. A holistic view of autonomy is achieved either by association of all of these requirements and associated enablers working together collaboratively or independently, and we now discuss these requirements in more detail.

Human role in autonomous systems

Even though a fully autonomous system that reaches the specified goal independently and without any human intervention is a design goal [28, 16, 92, 93, 9], many authors agree human intelligence and intervention remains key because of the safety, security, and bias issues posed by such systems [92, 93, 94, 95, 96, 97], and there is a need for keeping humans informed, enabling consent, and offering scope for intervention due to ethical reasons. Those issues have been quantified and reported in several publications in different fields to clarify the need for possible human intervention. In the context of surgical robots, 10,000 robotic-surgery-related adverse issues were recorded to the U.S. Food and Drug Administration (FDA) between 2011 and 2013, and the majority were due to malfunctions in the robot control system and instruments [98]. There have been major accidents related to autonomous cars due to them encountering new and ambiguous situations [98, 99].

While machines are good at tasks requiring fast computations, reacting quickly to known situations, and finding patterns in massive amounts of data, humans are good at resolving these new and ambiguous situations. Autonomous systems will require human support to guarantee correct, complete, and safe behaviour in all situations for the near future at least. Therefore, humans, machines and software must interact and collaborate to achieve the desired outcome [97, 100].

R. D. Alexander et al [101] defines the following possible human roles in autonomous systems:

1. **Human as a Safety Function**: human intervenes in the operation of the autonomous system when safety problems arise, acting as a safety function.

2. **Human does What Autonomous Systems Can't Do**: human performs tasks that cannot be automated due to technological limitations.

3. **Human as Necessary General Intelligence**: human performs tasks that require judgement, interpretation, and general human-level intelligence.

4. **Human as Mediator and Liaison**: human acts as a translator and interpreter between autonomous systems and other humans.

5. **Human as Maintainer**: human maintains the autonomous system when it is required.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

6. **Human as Peer**: human acts as a peer who cooperates with an autonomous system to achieve shared tasks.

7. **Human as Rescuer**: human rescues an autonomous system when it is immobilised or otherwise disrupted.

As illustrated by Figure 2.9, when the autonomy level of the system increases the role of human shifts from low-level and '4D' (dangerous, dirty, difficult, and dull) tasks to high-level and safe tasks [93, 33]. The human role can be broadly categorized into human-in-the-loop and human-on-the-loop systems in autonomous systems [93]. A human-in-the-loop system is a system where a machine executes a task while informing the worker, and may stop for human commands before continuing. A human-on-the-loop system is an autonomous system that executes the specified task independently and completely, and the role of a human in such systems is to monitor and to interfere if it is necessary. A high level of autonomy should not imply the exclusion of the human but allow for a seamless integration, both in the operational levels of controlling the process and to strategic levels of defining the aggregate plan, leading to higher levels of collaboration to achieve the common key performance indicators derived from manufacturing objectives [24, 102].

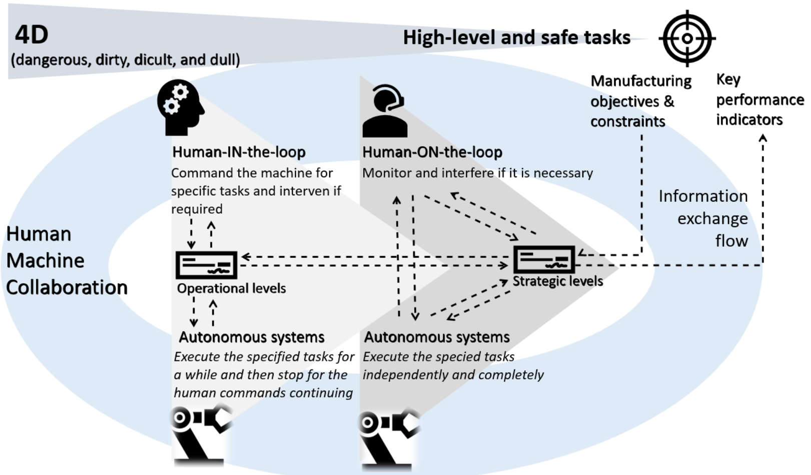

4D
(dangerous, dirty, dicult, and dull)
High-level and safe tasks
Manufacturing objectives & constraints
Key performance indicators
Human-IN-the-loop
Command the machine for specific tasks and interven if required
Human-ON-the-loop
Monitor and interfere if it is necessary
Information exchange flow
Human Machine Collaboration
Operational levels
Strategic levels
Autonomous systems
Execute the specified tasks for a while and then stop for the human commands continuing
Autonomous systems
Execute the specied tasks independently and completely

Figure 2.9: A description of possible human roles in autonomous systems

A substantial amount of work has been done in the context of human-in-the-loop systems [97, 103, 104, 105, 106, 51, 107, 108, 109]. [97] identified the technological challenges and limitations of integrating humans into the Cyber-Physical System autonomy loop. The Authors defined a conceptual framework that identifies the aspects that must be taken into consideration to design human participation, from both an abstract and engineering point of view. [103] developed a conceptual, technical road-map of autonomous pollination for future farming using micro-robotic air vehicle pollinators that are implemented using a combination of artificial intelligence and human expertise in-the-loop for the smart agricultural industry. [104] provides a theoretical framework and the conceptual instruments necessary for analyzing and understanding interactions with autonomous entities. The authors analyze the complex scenarios in which a cognitive agent has to decide if and how to delegate or adopt a task for another agent in a given context, how much autonomy is necessary for a given task, how trust and control play a relevant role in this decision, and how important their relationships and reciprocal influences are. [105] establishes an open prototyping platform and a design framework for rapid exploration of a novel human-in-the-loop application. The authors discuss the applications and challenges of human-in-the-loop Cyber-Physical systems. [106] proposes an approach to synthesize control protocols for autonomous systems that account for uncertainties and imperfections in interactions with human operators using Markov Decision Processes and stochastic two-player games. They demonstrate the applicability of the approach via a detailed unmanned aerial vehicle (UAV) mission planning case study. [51, 108] shows that human-in-the-loop robot teaching is effective in teaching robots to perform dynamic manipulation tasks in cooperation with a human partner; thus the robot simultaneously learns the action policy of the tutor and through time gains full autonomy.

Distributed Control Technologies for Autonomy

Traditional control architectures are based on hierarchical and centralized systems, focused on robustness and characterized by structural rigidity. This design approach gains in complexity as more tasks must be controlled i.e. higher numbers of interactions between the different components must be used to control the variety of functions executed [110]. In contrast, the objective of Autonomous Control is the realization of increased robustness and productivity of the entire system by correctly responding to dynamic and complex situations, but without increasing the overall system complexity. Machine autonomy is based on a processes of decision-making in heterarchical structures directly fed by raw data, digested, and turned into valuable information [111, 18].

The collection of this raw data is supported by the Internet of Thing (IoT) concept [4], and the related concept of the Industrial Internet of Things (IIoT). The goal is to avoid isolated systems based on proprietary solutions, and to enable data sharing by adopting common communication protocols that enable efficient and ubiquitous information aggregation and availability from the physical system [112]. Communication technologies can be divided into wired and wireless technologies, and different technologies are used to balance the following demands: i) high transmission rates, allowing large volumes of data to be shared; ii) high area coverage such as those encountered on a manufacturing shop floor; iii) low latency, to enable real-time applications (human-machine collaboration, remote real-time control); iv) large numbers of connections as the number of nodes grow exponentially; v) high reliability, especially for wireless technologies due to their high rate of package lost; and vi) high security, unsafe communication technologies may serve as an easy access point to malicious attacks and important data leakages [113]. However, the IIoT concept not only refers to the massive network of interconnected and heterogeneous devices (sensors, controllers, actuators, often referred to as operational technology), but also all the surrounding systems required to the management at upper layers (software data systems, often called information technology). High levels of abstraction try to describe a reference architecture for either IoT and IIoT deployment, based on their application context. A widely accepted three-tier pattern is presented in (Figure 2.10) [114].

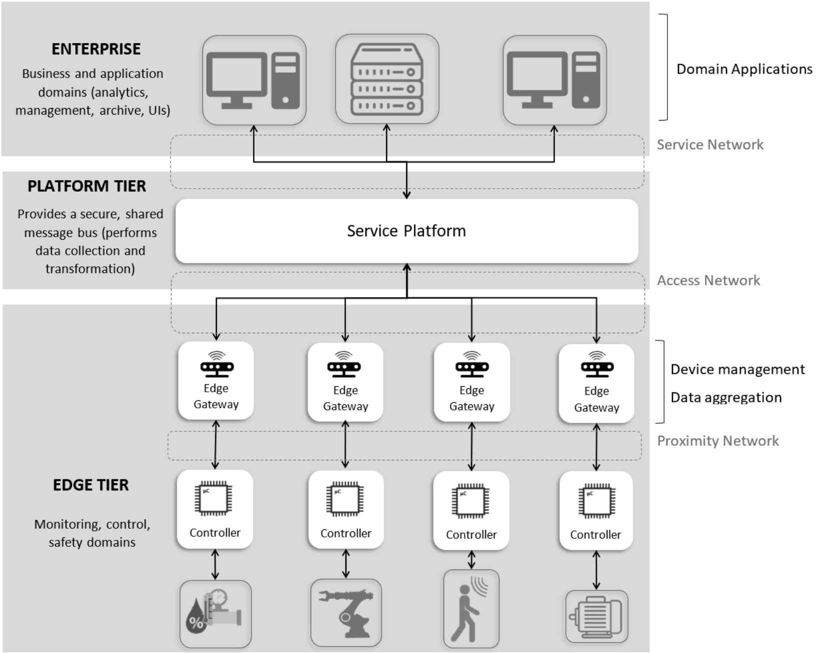


Figure 2.10: Three-tier IIoT hybrid architecture as a reference for IoT deployment in manufacturing to enable massive data gathering in autonomous systems. The proximity network collects data from edge nodes, forwarded over the access network to the platform tier, which processes the data for forwarding to the enterprise tier. The access network also processes and relays control commands from the enterprise tier back to the edge tier as well. The service network is used to communicate with the enterprise tier, which provides end users interfaces, control and domain-specific applications. Adapted from [114].

For better and effective decision-making, early studies highlighted the importance of knowledge representation, though the development from a single and unique model to multimodel-based strategies as industrial systems gain in complexity [33]. Models are explicit representations of the dynamics and responses of systems (linearly or non-linearly, statically or dynamically). Depending on the system being modelled and the objectives of the model, models are developed from first-principles understanding of the physics, from the captured

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

understanding of expert personnel, or directly from empirical data. The more complex a system is, the more dimensions and variables are involved and thus the more complex the model will be, or the larger the number of interconnected models required. The trend in industrial control is to increase the amounts of knowledge (in the form of models) to better understand system behaviour and to predict its future state. The improvements in dynamic resource management technology to allow multiple, competing, and heterogeneous computation tasks would permit the execution of multimodels on real-time platforms. In this context, Digital Twins (DT) are a core enabler of autonomy and key for achieving a new level of flexibility in Cyber-Physical Systems (CPS) [16], as the DT concept represents a standard framework or base for the deployment and real-time execution of different models [115].

DT was first presented as conceptual model for a virtual or digital representation equivalent of a physical product that includes the real space, virtual space, and the data and information flow between the two spaces. The interconnection and interoperability of the physical and the cyber world is a key aspect in its framework development and implementation for real-time operation control [50]. Since its first definition, DT properties evolved stating the following [116]:

1. It is at some level a replica of a real thing - the “physical twin”.
2. It is a digital representation stored on a computer.
3. It has a purpose of impacting an aspect of the environment in which its real counterpart exists, usually by serving one or more DT clients. It may have direct control of the physical twin, or act in an advisory role.
4. It uses models and simulations to achieve its purpose.
5. It incorporates some level of subject matter expertise in the solution. Some efforts illustrate that the combination of subject matter expertise and data provides more effective, robust and stable solutions in many manufacturing domains, than purely data-driven solutions.
6. It uses real-time data gathering to maintain the synchronization with its real counterpart.

Software and agent systems

Software systems may be implemented at all organizational and automation levels: from the lowest field levels of a shop floor, to the top managerial business levels [18, 117]. For instance, Autonomous Production Control (APC) methods used for control and resource management tasks based on different production scenarios [117, 118].

Together with DT models, another key technology for improving the performance in manufacturing systems is Artificial Intelligence (AI) [119]. AI applications based on machine learning (ML) are accepted as promising technologies in manufacturing, evolving the concept from ’Intelligent’ Manufacturing to

'Smart' Manufacturing which is knowledge-based and data-driven [120, 121]. ML techniques can be used to abstract high-level representations from large amounts of data, to assist decision making based on predictions – learning from those huge amounts of data. In this sense, learning processes may be classified as [122]:

1. **Supervised learning**: The correct response is provided by a teacher with experience of the problem.

2. **Reinforcement learning (RL)**: Only an evaluation of the chosen action is given by the teacher.

3. **Unsupervised learning**: No evaluation of the action is provided, since there is no teacher. The system must self-evaluate.

From a computation perspective, the implementation of smart systems is led by the use of Multi-Agent Systems (MAS) [123]. Each agent (which is a software entity), represents fundamental processing units [124], i.e. manufacturing resources, that build systems with cognitive capabilities and self-* behaviours such as self-awareness [125], self-organization [126]. These aspects enable as quick reactions to unpredictable perturbations [127] while increasing trustworthiness [128] and improving desirable properties in systems such as autonomy, responsiveness, redundancy, distributedness and openness. Centralized, hierarchical control architectures are replaced by loosely connected agents which focus on the control level, that not only offer a convenient way of controlling processes distributed over space and time, but as previous discussed enable the control of highly complex systems as a decentralized solution based on local decision making [129]. However, it is correspondingly more difficult to predict overall system behavior. MAS may be classified as either a) a centralized multi-agent coordination, in which a central agent undertakes the collection of partial plans, combines them and solves possible conflicts or b) a decentralized multi-agent coordination, in which agents communicate with each other for the creation of their plans and solution of possible arguments [50].

Autonomous behaviour can be implemented by swarms of cognitive agents (i.e. agents with AI techniques) [127] as well as other state-of-the-art design patters as well [131]. [119] proposes a framework that supports cognitive applications through linking a CPS with its DT in a closed-loop approach sharing both data and models. [130] provide a reinforcement learning-based intelligent agent (see Figure 2.11) which faces the problem of robotic control and human-robot interaction within manufacturing, presenting the potential for distributed agents to improve the adaptability of the system.

2.3.2 Existing Models and Level Definitions of Autonomy

References for the definition of autonomy levels in manufacturing

Research and standards in autonomous vehicles (e.g. self-driving cars) currently uses a useful - although imprecise - classification to distinguish levels

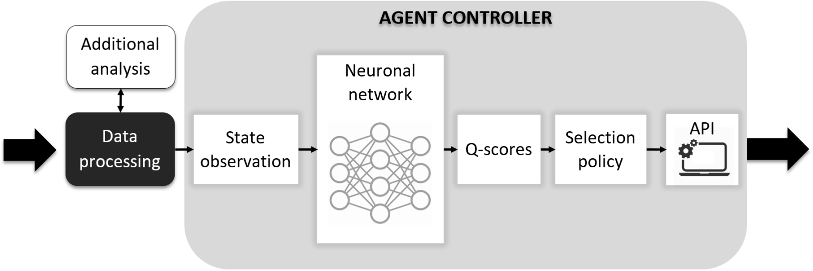
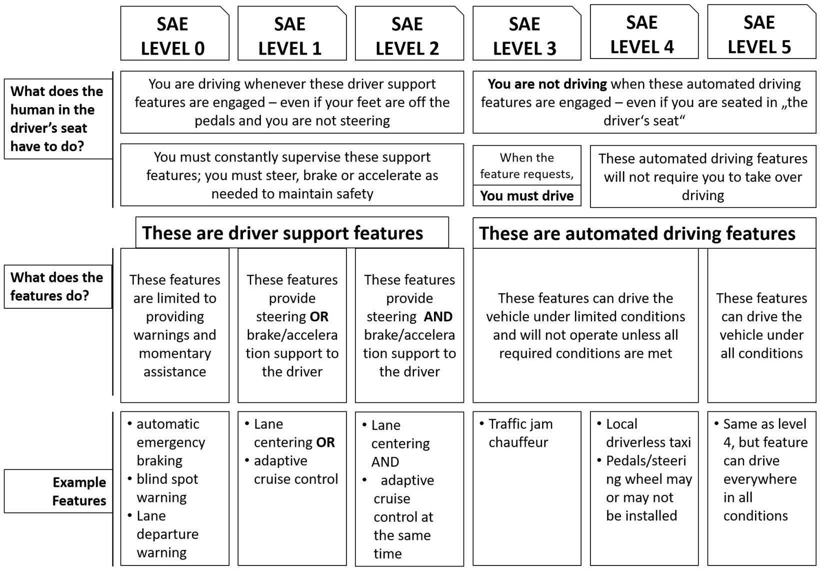


Figure 2.11: Reinforcement Learning (RL)-based Agent architecture for decision-making within a discrete action space. The learning agent provides adaptability and a autonomy via a control algorithm which generates instructions to adjust the operational parameters of robots in response to changes from its environment and collaborators. From a mathematical perspective, RL models a decision-making process based on a set of states, actions, and rewards. These actions are selected based on a policy which maximizes the cumulative sum of rewards as result of an optimum value function (Q-Score) which represents the quality of an action in a specific state. Adapted from [130].

of autonomy. There is a scale of five levels (plus a zero level) with level five used for “full autonomy” with no scope for human interaction, and level zero meaning no driver support systems at all. (Figure 2.12)

Compared to driving autonomy, Roland Berger developed the “Manufacturing Autonomous Level” (RB MAL) matrix to classify the level of autonomy in a manufacturing system, alsoon a scale from zero to five. [133] The RB MAL calculates this along the degree of automation (capability of production equipment to perform a defined set of tasks) and the degree of control intelligence (ability to orchestrate several production processes across multiple pieces of equipment)[133]. In terms of the degree of control intelligence, level zero involves completely manual and human-controlled operations, while level five includes mobile robots autonomously performing complex production and the use of production data to optimize schedules in real time. As for degree of automation (which is vertical axis of RB MAL) level zero describes completely manual operation, while level five includes flexible transportation of jigs, tools, and consumables.

Using this matrix, the study found that the semiconductor industry has the highest maturity levels, automotive suppliers are forward-thinking in terms of automation but lag behind in control intelligence, and the aerospace industry is still significantly behind[133]. By mapping out different industries on the RB MAL, the study identified key technical challenges facing autonomous manufacturing. The lower the needed external intervention by humans or other systems to achieve the goals under the uncertainties, the higher the degree of autonomy.



	SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?	You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			**You are not driving** when these automated driving features are engaged – even if you are seated in „the driver's seat"		
	You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, **You must drive**	These automated driving features will not require you to take over driving	
	These are driver support features			These are automated driving features		
What does the features do?	These features are limited to providing warnings and momentary assistance	These features provide steering **OR** brake/acceleration support to the driver	These features provide steering **AND** brake/acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met		These features can drive the vehicle under all conditions
Example Features	• automatic emergency braking • blind spot warning • Lane departure warning	• Lane centering **OR** • adaptive cruise control	• Lane centering AND • adaptive cruise control at the same time	• Traffic jam chauffeur	• Local driverless taxi • Pedals/steering wheel may or may not be installed	• Same as level 4, but feature can drive everywhere in all conditions

Figure 2.12: SAE J3016 Levels of driving automation [132]

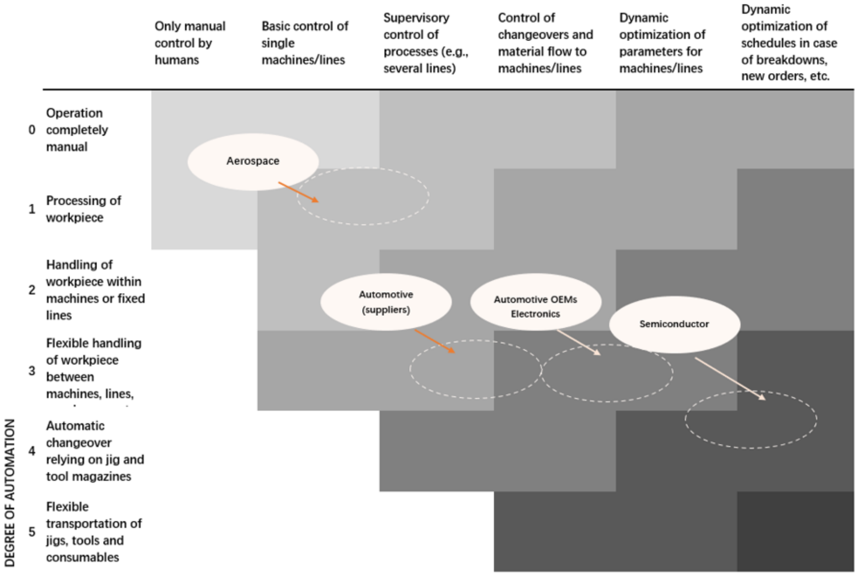

Only manual control by humans
Basic control of single machines/lines
Supervisory control of processes (e.g., several lines)
Control of changeovers and material flow to machines/lines
Dynamic optimization of parameters for machines/lines
Dynamic optimization of schedules in case of breakdowns, new orders, etc.
0 Operation completely manual
1 Processing of workpiece
2 Handling of workpiece within machines or fixed lines
3 Flexible handling of workpiece between machines, lines,
4 Automatic changeover relying on jig and tool magazines
5 Flexible transportation of jigs, tools and consumables
DEGREE OF AUTOMATION
Aerospace
Automotive (suppliers)
Automotive OEMs Electronics
Semiconductor

Figure 2.13: Roland Berger's Manufacturing Autonomy Levels [133]

Application

By mapping different factories into the RB MAL, it is possible to gain a clear picture of industry trends in autonomous production. While no industry has yet reached level five, there are clear leaders – mostly those with a short plant life-cycle and high volumes. The semiconductor industry has developed its own approach to transform manufacturing facilities into smart factories, and accordingly has the highest maturity levels. A significant number of automotive original equipment manufacture (OEM) factories also show mature degrees of automation and control intelligence. These companies drive innovation and invest more in fully automated projects and facilities, such as Mercedes' Factory 56; its new 'digital, flexible, green' production facility in Sindelfingen, Germany. [133]

Automotive suppliers are at a similar level to semiconductor fabrication in terms of automation, but lag behind when it comes to control intelligence. This is mainly because of their focus on the control of individual (customer) lines, instead of the autonomous control of the whole factory. As a result of complex structures and lower volumes, the aerospace industry is quite significantly behind. Large-scale investments in smart digital factories only occurred recently, with Turkish Aerospace Industries among the first to announce a smart factory in 2019.

Reference models

In terms of levels in manufacturing existing research has defined an model the connection between features and maturity levels. Features consist of manufacturing cell (automation, maintenance and autonomy), material information flow (automation of the material flow, agility of the material flow, connectedness of the information flow), development stages of entities (digitalization of entities, biologization of entities), Worker role (work integration), factory organization (cooperation between entities, versatile processes, structures and facilities, organizational form of the factory). [134] Five maturity levels are determined for each feature in an existing model. (Figure 2.14) Each line represent the corresponding maturity levels of the features in different stages of the whole manufacturing system.

- *Stage 0*: Analog Factory.

  This stage represents the initial stage in the model and contains all those factories in which neither automation nor digitalization have received any significant introduction so far. Essentially, workers take over the tasks that arise. Therefore, adaptable processes and a flexible material flow must be emphasized due to the lack of automation.

- *Stage 1*: Transparent Factory.

  In comparison to the previous stage, this stage is characterized by connected objects within the factory, an intelligent data connection of the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

equipment and a subsequent generation of information. The worker is thus always aware of the current status of the factory.

- *Stage 2*: Flexible Factory.

  Uses partially automated manufacturing cells. The information collected by decentralized networks can be analyzed and provides additional knowledge about the condition of the equipment. A modularized factory structure with distributed generation of information helps the factory to operate much more efficient and flexible.

- *Stage 3*: Semi-autonomous Factory.

  Automated manufacturing cells act independently on this stage. A comprehensive ICT network within the factory enables the systems to make decisions and creates changeable structures. The role of the worker primarily is to control and optimize manufacturing processes.

- *Stage 4*: Autonomous Factory.

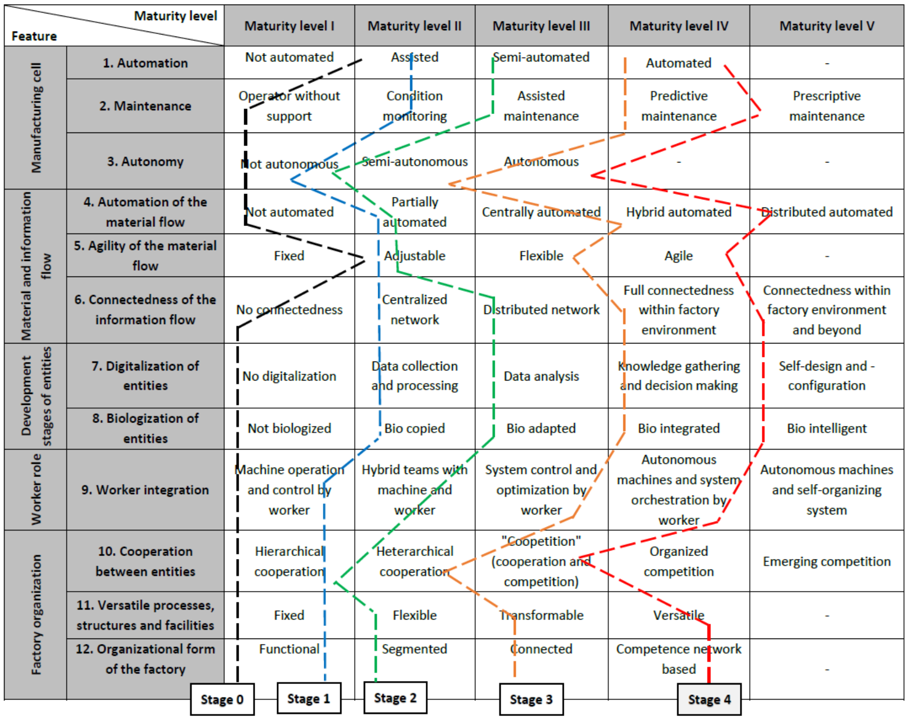
  With the final stage, the holistic autonomous manufacturing system is reached. The manufacturing facility, which is distinguished by its adaptability, requires ICT networking (also beyond the production environment) and autonomous sub-systems. The worker takes over the function of the orchestrator of the factory.

2.4 Defining Model of Autonomous Systems Requirements and Features

As presented in Section 2.2, the desired outcome of autonomy research and practice is a component that serves as the driver of independent intelligence in the distributed manufacturing system. However, there is a lack of standards and frameworks present to support whole-system autonomy, and this is particularly true for manufacturing autonomy. A high-level understanding of autonomy may exist, but with no way to break down autonomy into smaller components with clear requirements, implementation is a complex and difficult problem to solve.

The derived objectives and features from the literature review in this paper leads to our definition of ‘Autonomy in Manufacturing’. This is not a short textual definition as one might find in a dictionary, but a comprehensive model of the features and requirements of autonomy for distributed manufacturing systems intended to facilitate research and implementation in the domain. This process also describes the incremental steps to achieve manufacturing autonomy. In order to have a better understanding of the autonomy levels of the manufacturing levels, the deduction of features and functionalities to describe an autonomous manufacturing system is primarily based on the analysis of existing concepts for distributed manufacturing systems (see Table 2.2).This

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078



Feature \ Maturity level		Maturity level I	Maturity level II	Maturity level III	Maturity level IV	Maturity level V
Manufacturing cell	1. Automation	Not automated	Assisted	Semi-automated	Automated	-
	2. Maintenance	Operator without support	Condition monitoring	Assisted maintenance	Predictive maintenance	Prescriptive maintenance
	3. Autonomy	Not autonomous	Semi-autonomous	Autonomous	-	-
Material and information flow	4. Automation of the material flow	Not automated	Partially automated	Centrally automated	Hybrid automated	Distributed automated
	5. Agility of the material flow	Fixed	Adjustable	Flexible	Agile	-
	6. Connectedness of the information flow	No connectedness	Centralized network	Distributed network	Full connectedness within factory environment	Connectedness within factory environment and beyond
Development stages of entities	7. Digitalization of entities	No digitalization	Data collection and processing	Data analysis	Knowledge gathering and decision making	Self-design and -configuration
	8. Biologization of entities	Not biologized	Bio copied	Bio adapted	Bio integrated	Bio intelligent
Worker role	9. Worker integration	Machine operation and control by worker	Hybrid teams with machine and worker	System control and optimization by worker	Autonomous machines and system orchestration by worker	Autonomous machines and self-organizing system
Factory organization	10. Cooperation between entities	Hierarchical cooperation	Heterarchical cooperation	"Coopetition" (cooperation and competition)	Organized competition	Emerging competition
	11. Versatile processes, structures and facilities	Fixed	Flexible	Transformable	Versatile	-
	12. Organizational form of the factory	Functional	Segmented	Connected	Competence network based	-

Figure 2.14: Features and maturity levels of Autonomous Production [134]

is achieved using a process called a Morphological Box, which is designed for breaking complex concepts into quantifiable sub-components [135, 136, 137].

The authors present a five-level model of autonomy to describe and characterize different levels of whole-system autonomy and the requirements for each, to enable evaluation of an entire manufacturing system. We also determine three to five levels of technological maturity for each functionality, and then describe them. Each feature maturity level is then matched to a whole-system autonomous level, enabling the identification and evaluation of manufacturing system's autonomy capability based on the level of technological maturity of the enabling features.

- *Autonomous Level 1 (AL1)*: Factory without autonomy.

  The first level covers manufacturing systems with the lowest autonomy level even if the system contains high levels of automation. Mainly, the systems relies on operator actions to start/continue processes, and no highly connected systems are proposed here.

- *Autonomous Level 2 (AL2)*: Basic Connected Factory.

  An initial attempt to address connectivity in manufacturing. This is defined by the inclusion of centralised data-driven management in connected systems, to allow improved functionality and context-aware features. Human operators are still needed to intervene for many tasks, particularly with high uncertainty or failure-recovery processes.

- *Autonomous Level 3 (AL3)*: Distributed Intelligent Factory.

  This level represents the distributed manufacturing intelligence concept, and is characterized by the introduction of predictive features and self-adaptable behaviors to external inputs locally to assets. Human operators receive suggestions for alternative or optimized activities but the main system objectives are still under the human control during operations as the work space is fully shared. Collected data is continuously monitored and automatically managed ready to be shared between all connected systems.

- *Autonomous Level 4 (AL4)*: Semi-Autonomous Factory.

  At this stage, the system determines actions based on its high context-awareness and its intrinsic safety rules. Human operators work cooperatively with the system rather than directly instructing it. Even if the system is able to analyze its environment and track states for 'business as usual' execution, system goals and response to major disturbances are still monitored by humans.

- *Autonomous Level 5 (AL5)*: Fully Autonomous Factory.

  The system is fully self-adaptable to uncertain or unforeseen inputs, enabled by advanced self-learning capabilities. The system is able to decide

Table 2.2: Morphological box for the Autonomy concept and Five-Stage model features

Category	**Feature**
Data, information and knowledge	Data Knowledge management Interoperability
Process	Synchronization Functionality Optimization Reliability
Interaction	Context Awareness Interaction with humans Safety
Infrastructure	Connectivity Industry Control Cyber-security
Self-X	Self-configuration Self-healing Self-optimization Self-protection
Measurement performance	Accuracy Stability Certainty

by itself the best option to meet common goals for these connected manufacturing systems without human intervention. However, AI and humans will still collaborate actively during manufacturing operations.

2.4.1 Data, information and knowledge

This category groups the features: data, knowledge management and interoperability, with their respective functionalities (Table 2.3).

Table 2.3: Features and functionalities of data, information and knowledge

Category	Feature	Functionality
Data, information and knowledge	Data	Data structures Data processing Data integration
	Knowledge management	Knowledge management
	Interoperability	Hardware, software and knowledge

Data

The data feature can be divided into three functionalities: data structure and data processing (both with three maturity levels); and data integration with five. This is depicted in Fig. 2.15. Data structures develop in maturity as they progress from unstructured to full structured data types, and data processing matures as it transitions from data collection to data analysis that allows real-time knowledge processing [138]. For data integration to reach a higher level of maturity, it has to go through the following maturity levels [139]:

1. At this stage IT integration is non-existing, data about the manufacturing process is neither stored nor used in any way, and manufacturing equipment like machines, tools and work-parts are not integrated with the IT systems;

2. The traditional information pyramid of manufacturing is implemented, machines are integrated and managed by a Manufacturing Execution System (MES);

3. Relevant manufacturing data is integrated with data from other departments;

4. The implementation of a Service-Orientated Architectures allows data provisioning. To eliminate inefficient communication, an enterprise service bus connects data between enterprise and shop floor systems;

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

5. Advanced real-time analytics play a major role for extracting information from data, therefore bringing the need for digital twin (or other system) which integrates all systems, devices, and data across the entire product life cycle and uses insights about this data to automatically optimize the factory and all manufacturing processes.

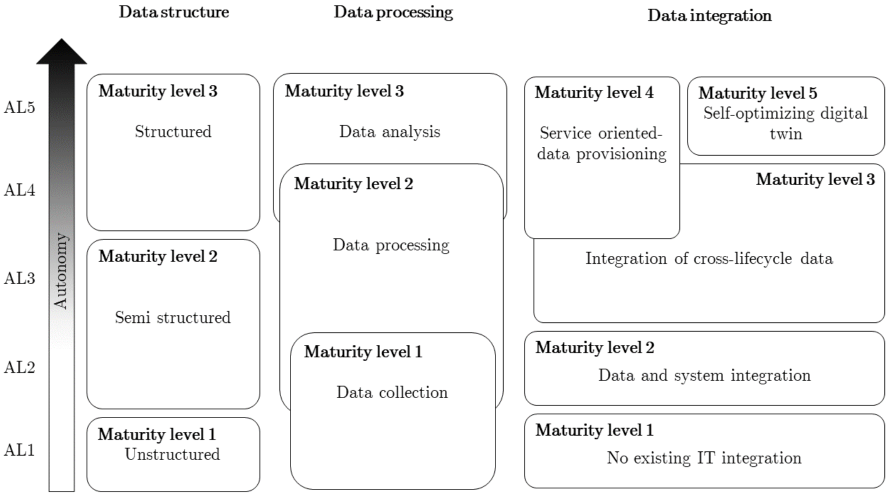


Figure 2.15: Maturity levels of data functionalities

Knowledge management

Autonomy also requires knowledge management—the ability to manage and deploy acquired knowledge effectively. The lowest level of autonomy does not require automated knowledge management, but as the level of autonomy increases knowledge management becomes a crucial part of an autonomous system. Knowledge management is done by explicit knowledge design and analyzing historical data.

However, the designed knowledge may not be optimal for the system. Therefore, as the technology progresses to the most advanced maturity level of knowledge management will shift from explicit hand-design to automatic making sense of underlying processes by analyzing raw data (Fig. 2.16):

1. No knowledge management is utilised;

2. Low-level knowledge design of entities and running processes that help identify the current context;

3. Explicit knowledge design is required for the system to be context-aware and self-adaptable to disturbances, though human operator feedback is considered;

4. Knowledge management is a combination of explicit knowledge design and historical data, so the system self-improves with every iteration of the design;

5. No explicit knowledge design is required by humans, the system is able to analyze the raw data of its underlying processes.

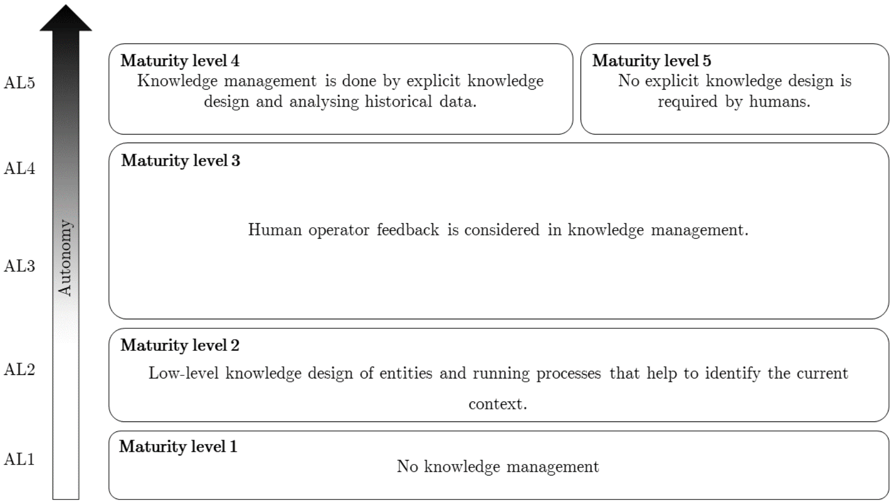

AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 4
Knowledge management is done by explicit knowledge design and analysing historical data.
Maturity level 5
No explicit knowledge design is required by humans.
Maturity level 3
Human operator feedback is considered in knowledge management.
Maturity level 2
Low-level knowledge design of entities and running processes that help to identify the current context.
Maturity level 1
No knowledge management

Figure 2.16: Maturity levels of knowledge management

Interoperability

As communication between the system's components is a key enabler to achieving a high level of autonomy, from the data point of view interoperability must be considered as one of its prerequisites. Information sharing would be necessary between all business levels, and therefore is not only limited to software or hardware infrastructure, but to the process and information itself, supported by widely spread standards and validation methods.

At the lower levels of maturity (Fig. 2.17) , interoperability is limited to compatibility at the infrastructure level. As the maturity level increases, standardization at production and business-related processes and information management is required, followed by protocols and methods to guarantee knowledge interoperability. At the top levels of maturity, methods for certification and validation of interoperability would be required, combined with interoperability of non-technical high level goals and business priorities [140, 141].

2.4.2 Process

The category encapsulates the features and functionalities of synchronisation, functionality, optimization and reliability (Table 2.4).This category relates to

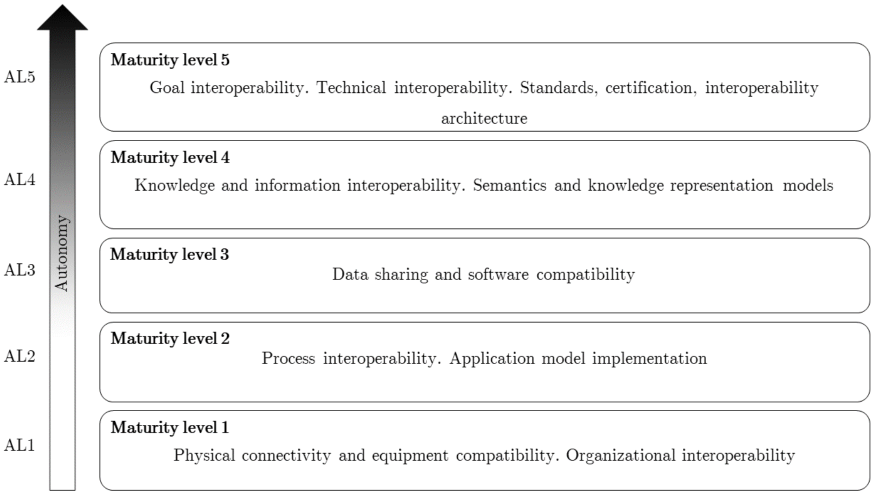

AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 5
Goal interoperability. Technical interoperability. Standards, certification, interoperability architecture
Maturity level 4
Knowledge and information interoperability. Semantics and knowledge representation models
Maturity level 3
Data sharing and software compatibility
Maturity level 2
Process interoperability. Application model implementation
Maturity level 1
Physical connectivity and equipment compatibility. Organizational interoperability

Figure 2.17: Maturity levels of interoperability

features that form or influence a process of a production environment. End to end synchronisation along with implicit functionality of a process direct towards autonomous behavior. This behavior is made consistent by process optimization and reliable execution.

Table 2.4: Features and functionalities of Autonomous Processes

Category	**Feature**	**Functionality**
Process	Synchronisation	Material/logistics Physical (Production Machines) [142]
	Functionality	Mapping from Input (X) to Output (Y)
	Optimization	Optimization function Productive and Unproductive contribution
	Reliability	Failure recovery

Synchronisation

The synchronization aspect in production system autonomy can be categorized into material/logistic synchronization, and physical synchronization as per the work done by [142]. The material/logistic synchronization is concerned with part and material movement from one station to ensure stations are not left idle, and physical synchronisation is concerned with regular controlled part processing at each station. At the lowest maturity levels there exists no logistic synchronisation between each station. The part/material movement follows a manual philosophy with automation increasing with each autonomy level. At level three physical synchronisation is realized where parts are produced at a synchronous rate at each station and transported regularly to match that physical synchronous effect. At level four the system could maintain this behaviour independently but still requires human-assistance to start, stop or change certain aspects of behaviour. At level five this becomes completely independent without human-assistance i.e. the system decides the rate and other conditions itself for synchronization.

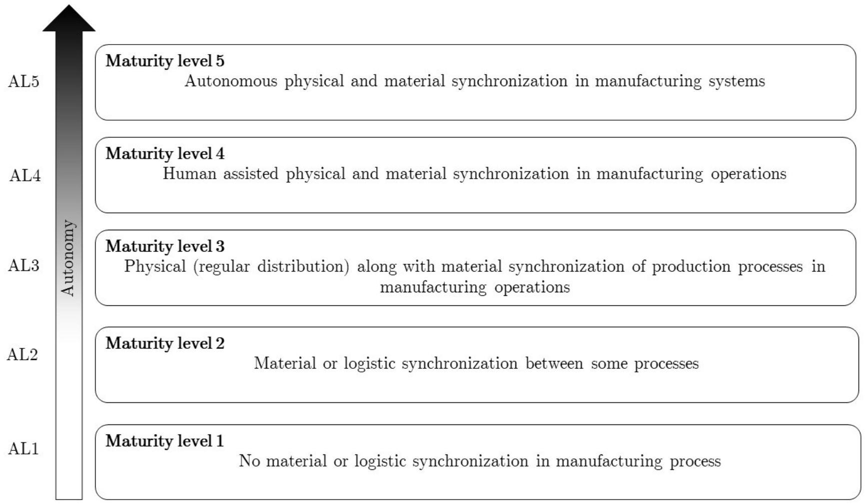

AL5
Maturity level 5
Autonomous physical and material synchronization in manufacturing systems
AL4
Maturity level 4
Human assisted physical and material synchronization in manufacturing operations
AL3
Maturity level 3
Physical (regular distribution) along with material synchronization of production processes in manufacturing operations
AL2
Maturity level 2
Material or logistic synchronization between some processes
AL1
Maturity level 1
No material or logistic synchronization in manufacturing process
Autonomy

Figure 2.18: Maturity levels of synchronisation

Functionality

Functionality in a production system maps an input to output. A function can be defined as an implicit event wherein the consecutive output is a function of the input and the previous output. In manufacturing context, this concept is taken as conversion of input to output by a production system element. At the basic level of autonomy the production system is simply converting input to output. At this stage however,a system is not capable of associating to multiple inputs and multiple outputs. So, in true essence a true implicit function is not possible. As the autonomy level increases such association is realized. Further up the autonomy scale inputs can be automated and machine programmed to produce output functionality by mapping outputs to inputs.In the level-4 production system looks to the assigned goal to develop suitable input-output ordered pairs. Execution of functionality still is retained by selection of human operator. In case of full-autonomy the human-operator does not get involved in any selection process and completed mapping and execution is performed independently as per pertaining goals.

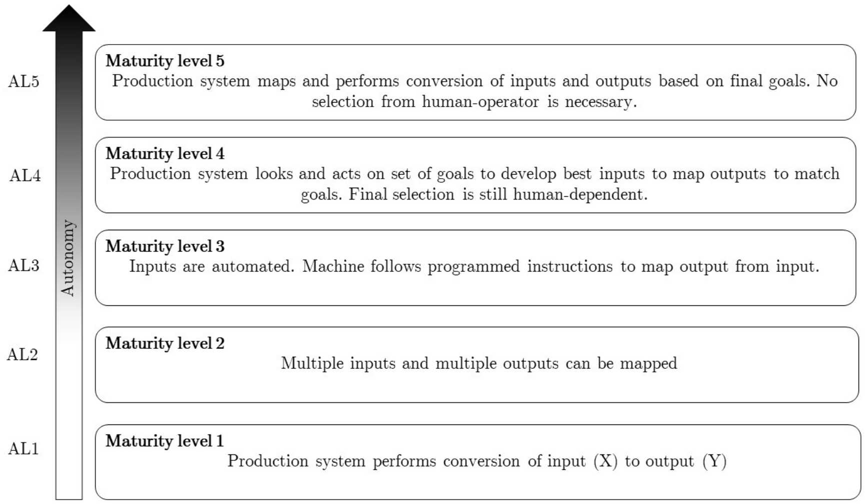
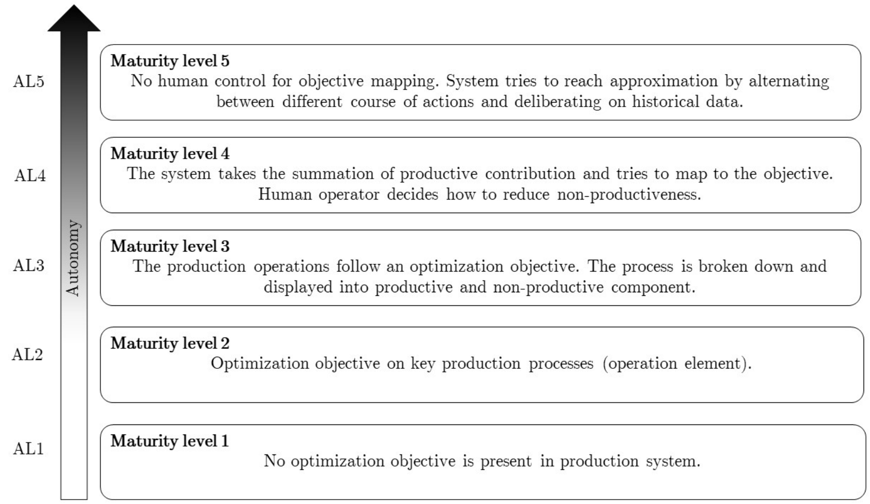

AL5
Maturity level 5
Production system maps and performs conversion of inputs and outputs based on final goals. No selection from human-operator is necessary.
AL4
Maturity level 4
Production system looks and acts on set of goals to develop best inputs to map outputs to match goals. Final selection is still human-dependent.
AL3
Maturity level 3
Inputs are automated. Machine follows programmed instructions to map output from input.
Autonomy
AL2
Maturity level 2
Multiple inputs and multiple outputs can be mapped
AL1
Maturity level 1
Production system performs conversion of input (X) to output (Y)

Figure 2.19: Maturity levels of functionality

Optimisation

Optimisation deals with identifying productive and non-productive contribution of the operational elements of a system, and the influencing factors. At the basic maturity level there is no capability in the production system for optimisation. As the autonomy level increases the capability is present at key production processes (operation elements). At a higher level the whole production line is assumed to collectively follow an optimization objective. Moreover, at this level it is possible to assess and display the productive and non-productive components of the operation. Level four takes a summation of all productive contributions and links it to the objective i.e. both the total productive and non-productive contributions are identified, but a human-operator is required to decide on how to reduce non-productiveness. At level five no human control for objective mapping is necessary and the system reaches best approximation by itself.



Maturity level 5
No human control for objective mapping. System tries to reach approximation by alternating between different course of actions and deliberating on historical data.

Maturity level 4
The system takes the summation of productive contribution and tries to map to the objective. Human operator decides how to reduce non-productiveness.

Maturity level 3
The production operations follow an optimization objective. The process is broken down and displayed into productive and non-productive component.

Maturity level 2
Optimization objective on key production processes (operation element).

Maturity level 1
No optimization objective is present in production system.

Figure 2.20: Maturity levels of optimisation

Reliability

The reliability of the system is another aspect to consider when thinking about the autonomous functioning of a system. The ability to overcome failures during operation is a critical feature that would ensure continuous and effective operation. At maturity level zero a system has no ability to detect or respond to failures. As maturity increases the system gains the ability to diagnose failures and gather the information that allows the system to predict them. Going further, higher levels of maturity will have a recovery routine that will help the system to reallocate resources to avoid stops or disruption to its operation, which could be done at software and hardware level (if backup systems are included). This may involve informing human workers on what actions are required or (at level five) executing the recovery plan itself.

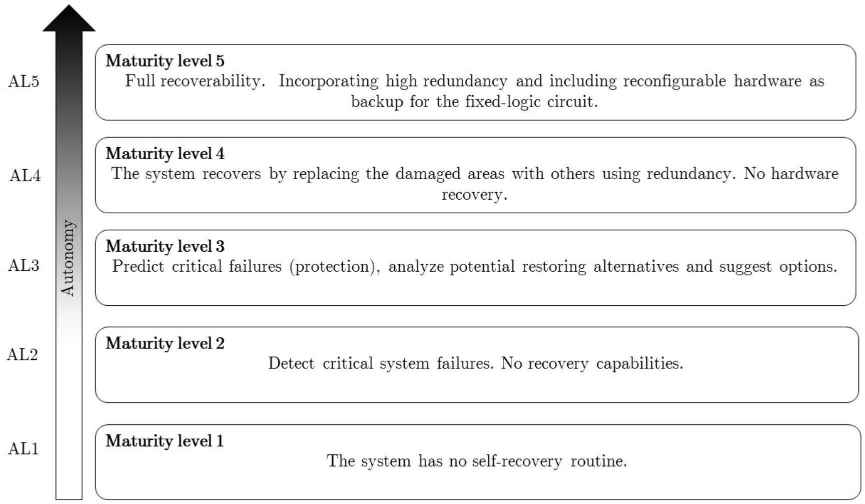

AL5
Maturity level 5
Full recoverability. Incorporating high redundancy and including reconfigurable hardware as backup for the fixed-logic circuit.
AL4
Maturity level 4
The system recovers by replacing the damaged areas with others using redundancy. No hardware recovery.
AL3
Maturity level 3
Predict critical failures (protection), analyze potential restoring alternatives and suggest options.
AL2
Maturity level 2
Detect critical system failures. No recovery capabilities.
AL1
Maturity level 1
The system has no self-recovery routine.
Autonomy

Figure 2.21: Maturity levels of reliability

2.4.3 Interaction

Seamless interaction among humans, machines and smart devices constitutes a distinctive requirement of autonomous manufacturing systems that are able to control the inter-connected manufacturing process flows. Table 2.5 indicates autonomous features enabling system interactions, including context aware, human interaction and safety.

Table 2.5: Features and functionalities of different interactions

Category	**Feature**	**Functionality**
Interactions	Context aware	Physical, virtual and user environment
	With humans	Operation Programming Control
	Safety	Integration level

Context-aware Interaction

As ICT tools in the factory are increasingly used as part of the digital and smart manufacturing paradigms, controlling access to data and understanding the required information flows poses challenges to be dealt with. Those challenges require context-awareness in autonomous manufacturing systems [143, 144, 145]. Context awareness, based on the information obtained from cyber-physical systems, is characterized as as the ability of the system to sense, interpret, adapt, discover contextual resources and associate digital data with the user's context to better understand how data should be used [146, 147].

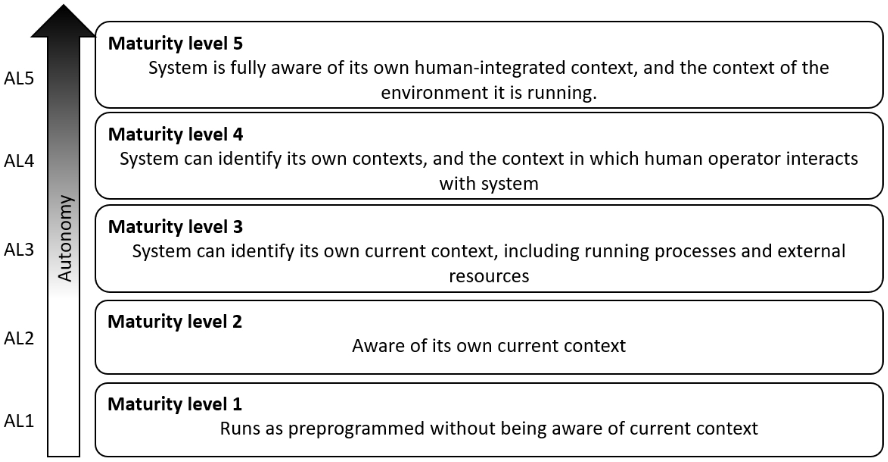

AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 5
System is fully aware of its own human-integrated context, and the context of the environment it is running.
Maturity level 4
System can identify its own contexts, and the context in which human operator interacts with system
Maturity level 3
System can identify its own current context, including running processes and external resources
Maturity level 2
Aware of its own current context
Maturity level 1
Runs as preprogrammed without being aware of current context

Figure 2.22: Maturity levels of context awareness

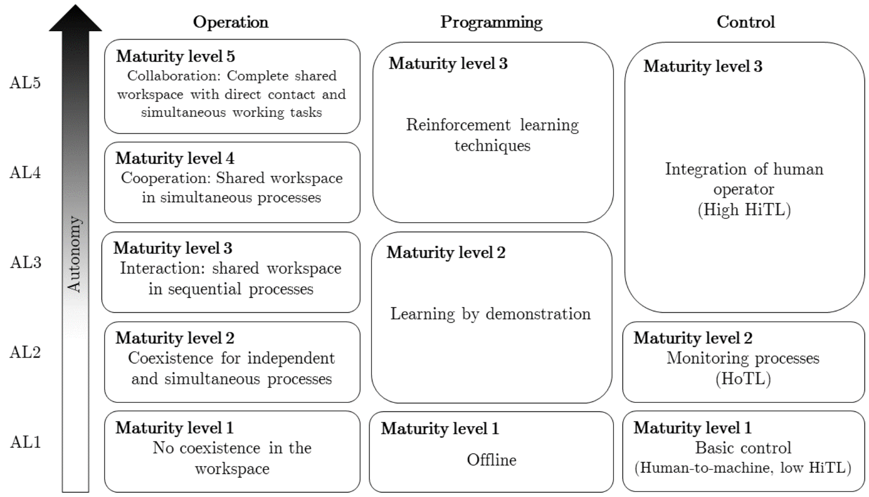
The increase in the level of autonomy will require the enhancement of context-awareness feature as per Figure 2.22. Non-autonomous systems run pre-defined programs without being aware of surrounding contextual information and integration. As maturity increases, systems become aware of its current context and adapt their operations to running processes and external resources. At the highest maturity levels, the system can identify its current context, running processes, external resources and the context in which human operator interacts with a system and context of the environment, and thus aim at increasing usability and effectiveness by taking environmental context into account.

Interaction with humans

The interaction between humans and robots has been widely studied during the last years from different viewpoints [148]. However, this general classification of the relationship between both actors may be extrapolated to autonomous systems as well (see Figure 2.23), as it is analysed under a perspective of the autonomy concept. This analysis requires special effort since the future of human-machine collaborations must be an attempt to enable future workforce to handle complexity by complementing and enhancing rather than replacing human capabilities and skills [149].

From an *operational* view [148], a basic level requires safety distancing and protective measures as the system is not able to contextualise interactions between both actors. As maturity increases, this permits the coexistence of both actors without overlapping each other, neither direct contact and the processes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078



	Operation	Programming	Control
AL5	**Maturity level 5** Collaboration: Complete shared workspace with direct contact and simultaneous working tasks	**Maturity level 3** Reinforcement learning techniques	**Maturity level 3** Integration of human operator (High HiTL)
AL4	**Maturity level 4** Cooperation: Shared workspace in simultaneous processes		
AL3	**Maturity level 3** Interaction: shared workspace in sequential processes	**Maturity level 2** Learning by demonstration	
AL2	**Maturity level 2** Coexistence for independent and simultaneous processes		**Maturity level 2** Monitoring processes (HoTL)
AL1	**Maturity level 1** No coexistence in the workspace	**Maturity level 1** Offline	**Maturity level 1** Basic control (Human-to-machine, low HiTL)

Figure 2.23: Maturity levels of human interactions

are performed independently and simultaneously. Shared work spaces refers to human and the system working in the same working area with no physical or virtual fencing for separation. In that sense, a more interactive level allows both actors to share the work space and communication between each other is required. At this level, the work space is shared and physical contact is allowed. A cooperative operation is developed when both actors have their own objectives and goals in a mutually beneficial perspective, but normally they must wait the other's availability to work simultaneously. Here physical, computational and cognitive resources may be shared temporarily. The highest level of maturity allows a complete collaboration between actors. Both actors have compatible objectives and goals and the work follows coordinated and synchronised operations in a fluent manner.

Offline programming is mainly based on writing instructions and directly uploading them to the system. This requires human operators with a high level of knowledge of programming method, the task, and the system capabilities. Higher maturity in this areas enables learning through demonstration, based on repetition of movements previously guided by the human operator. It also requires a human operator with a high level of knowledge of the processes but the program is written directly by the system (storing positions, velocities and other parameters) so requires less knowledge of the method of programming. The highest level of autonomy for programming is when the system autonomously knows and understands its own capabilities, as well the tasks it has to perform without specific human guidance. [150].

As defined in previous sections, Human-*in*-the-loop (HitL) *control* allows the integration of a human operator in the system. As almost any system requires some minimal human intervention (system start / shutdown, continue command, etc), a low level of mono-directional (i.e. human to system) interaction is required. However, at higher autonomy levels, the control role over the system can move from a more passive one (Human-*on*-the-loop, only for monitoring and correcting purposes when necessary) to a more active one. In this last case (high HitL), the human operator is an integrated element of the system during main operations and tasks execution, with bi-directional information flow and awareness between the system and the operator.

Safety

Regarding how safety features are deployed in autonomous systems, we mainly classify them as intrinsic or extrinsic (see Figure 2.24). The level of uncertainty is then analysed depending on the deployment level, as it plays a major role [151].

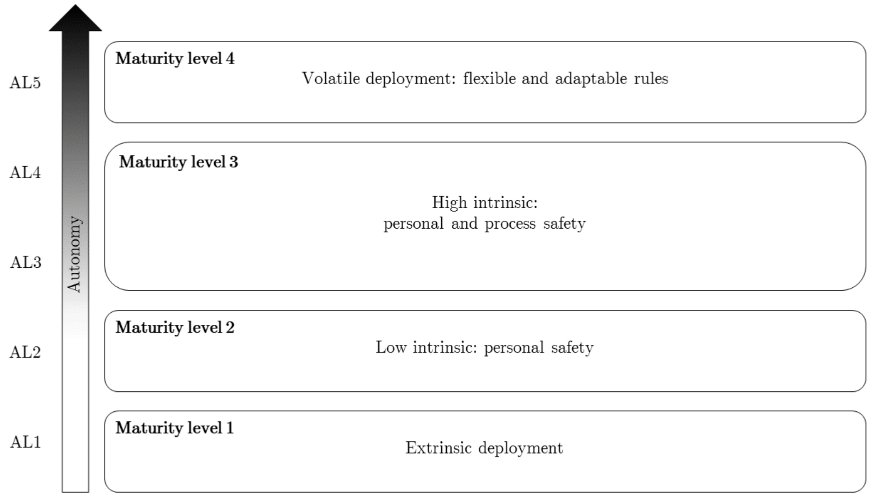


Figure 2.24: Maturity levels of different integration levels of safety features

Low levels of safety maturity use extrinsic safety measures, which are external to the system and built by complementary components or systems such as light gates or floor scanners. At this level, the system is built rigidly and uncertainty is reduced at its minimum. On the other hand, medium maturity can use intrinsic safety measures composed of built-in rules or components. Systems consider safety measures before acting and require the participation of workers in rule management. A higher level of maturity will assure process safety as well as personal one and, in addition, certain level of rules adaptation is required. The highest maturity level includes flexible rules with high levels of adaptation to enable the system to respond safely and predictably to dynamic, changing and novel scenarios while at the same time being called upon to plan and reflect on their actions.

2.4.4 Infrastructure

Ubiquitous connectivity infrastructure is a requirement in realizing Industry 4.0. With a rise in reliance on cloud technologies, One of the major pitfalls—even in developed nations such as Germany—is a lack of reliable high-speed broadband for SMEs. [152]. Industry 4.0 would require every supply chain member to be integrated; thus, digital infrastructure is a factor that cannot be ignored. The results of the survey of Penton's IoT Institute, wherein 33 % of respondents believed that lack of infrastructure is an issue in adopting IoT. The survey further found that many firms are now collaborating instead of competing regarding the necessary infrastructure development needed for Industry 4.0.[153]

Table 2.6: Features and functionalities of infrastructure

Category	Feature	Functionality
Infrastructure	Connectivity	Types of network Application levels
	Industry Control	Control Technology development Application levels
	Cybersecurity	Hardware, software and knowledge

Connectivity

The degree of connectivity is divided into two subcategories: the type of network and the application level. The type of network will influence the available data exchange strategies and efficiencies, while also determining how agile the system is to change or scaling. The application level is the elements of the production system, from the lowest process and field levels, to the highest connectivity between geographically distributed locations. The efficiency of a system depends on how well the individual elements of the production system are networked, as this is a requirement for data exchange and collaboration in autonomous systems.[154]

In addition to the degree in which locations are networked, the type of network is also an important part of connectivity. At the lowest maturity level system elements are not networked, or rely on a manual movement of data. At maturity level two centrally networked elements receive both information and feedback via a central unit. Communication with other systems takes place via the central unit, so there is no direct communication between the elements. At maturity four decentralized networked production elements can communicate with each other without a central unit, which facilitates robustness to failure

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

and simpler addition or removal of elements and systems. Between centralised and decentralised, a hybrid system communicates partly in a decentralized manner and partly via the central unit. Here, some of the elements can already communicate with each other but the main communication takes place via the central unit.

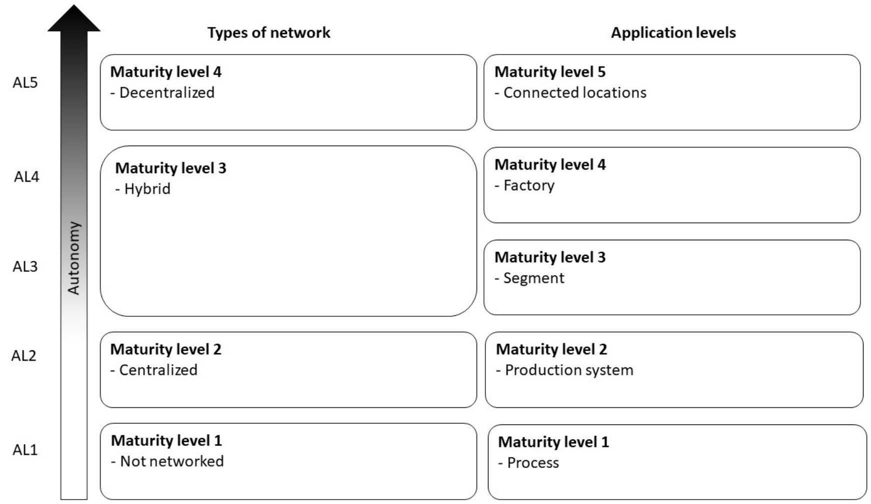


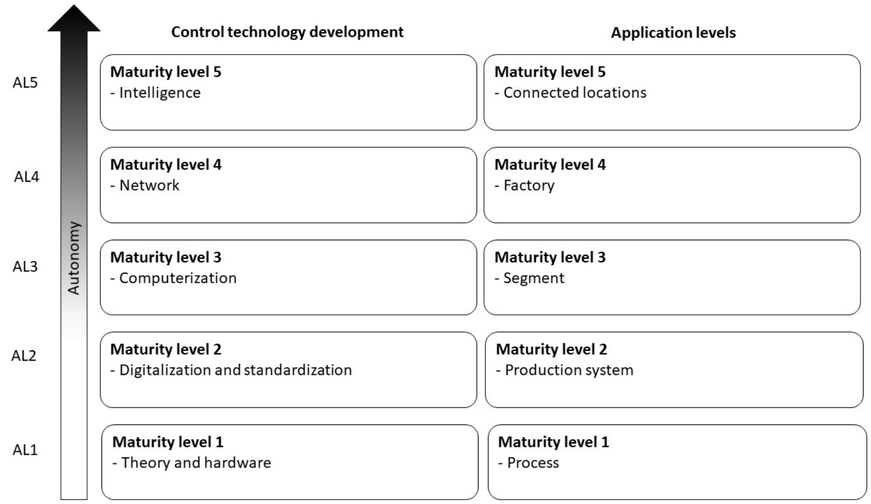
Figure 2.25: maturity levels of connectivity

Industry control

Industrial control refers to industrial automation or operational technology, which use the combination of electronic and electrical, mechanical, software, computer technology, and microelectronics technology to control manufacturing processes to achieve their goals. Industrial control aims to make the production and manufacturing processes more automatic, efficient, accurate, and controllable and visible. For the development of industry control technology, two aspects and five maturity levels for each aspect are categorised. The Two aspects are are control technology development and applications levels. For control technology development, at maturity level 1, the theory and hardware are developed. While at maturity level 5, the intelligence technology for control theory should be developed. For Application levels, at maturity level 1, the industry control is only applied in process.As the applications level goes up, the industry control can be used in bigger and more places.At maturity level 5, it will be expanded to connected locations.

Cybersecurity

With the generation of valuable information, there arises a need to secure proprietary data and processes from malicious actors. Cybersecurity deals with



	Control technology development	Application levels
AL5	Maturity level 5 - Intelligence	Maturity level 5 - Connected locations
AL4	Maturity level 4 - Network	Maturity level 4 - Factory
AL3	Maturity level 3 - Computerization	Maturity level 3 - Segment
AL2	Maturity level 2 - Digitalization and standardization	Maturity level 2 - Production system
AL1	Maturity level 1 - Theory and hardware	Maturity level 1 - Process

Figure 2.26: maturity levels of industry control

the protection of the whole information technology and operational technology infrastructure. At maturity one, the minimum required security level is implemented (i.e. mandated by government safety standards). With the increase of the maturity level to two, infrastructure is secured through the continuous monitoring of the running processes and connected devices. At maturity three possible attack vectors can be predicted and reported to a human operator. The highest maturity level four has automatic intrusion detection directly implemented into computers or controllers, or separate hardware with higher computational power and system can secure itself from most of the attacks [155].

2.4.5 Self-x capabilities

Self-x capabilities in manufacturing take roots from the concept of Autonomic Computing in IT industry. In the 2001 manifesto of IBM, Horn [156] proposes a possible solution for the growing software complexity crisis. The proposed approach to solving the challenge is an Autonomic Computing system which has self-management abilities similar to the human autonomic nervous system. Autonomic Computing systems are capable of running themselves, adjust to varying circumstances, and handle their resources efficiently. Self-management abilities of Autonomic Computing systems comprise of four major characteristics: self-configuration, self-healing, self-optimization and self-protection. In literature, these four main characteristics of Autonomic Computing are known as self-CHOP capabilities:

- **Self-configuration** is the ability of a system to configure and reconfigure itself under changing and unpredictable conditions [156, 157, 158],

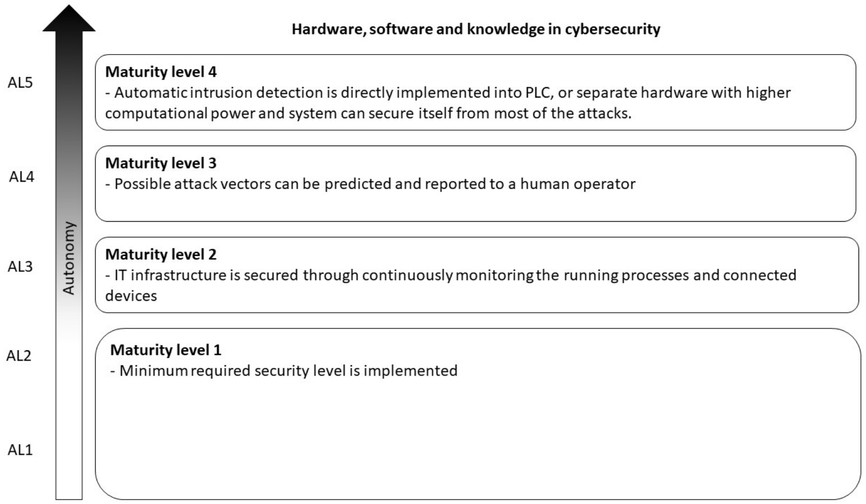

Hardware, software and knowledge in cybersecurity
AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 4
- Automatic intrusion detection is directly implemented into PLC, or separate hardware with higher computational power and system can secure itself from most of the attacks.
Maturity level 3
- Possible attack vectors can be predicted and reported to a human operator
Maturity level 2
- IT infrastructure is secured through continuously monitoring the running processes and connected devices
Maturity level 1
- Minimum required security level is implemented

Figure 2.27: Maturity levels of cybersecurity

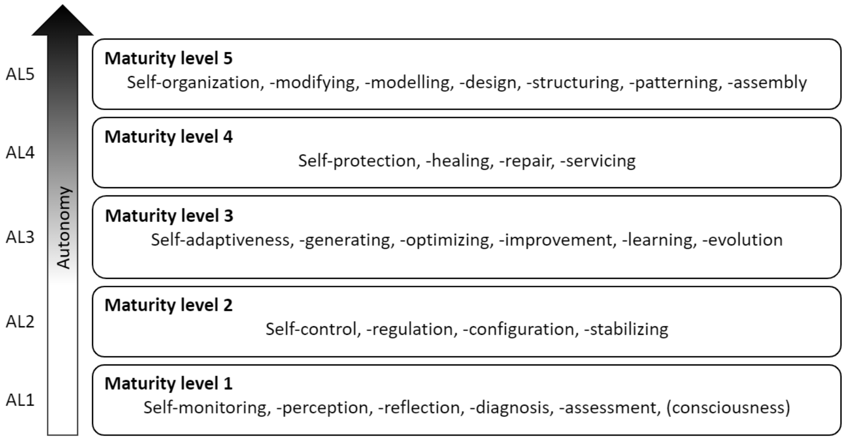

AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 5
Self-organization, -modifying, -modelling, -design, -structuring, -patterning, -assembly
Maturity level 4
Self-protection, -healing, -repair, -servicing
Maturity level 3
Self-adaptiveness, -generating, -optimizing, -improvement, -learning, -evolution
Maturity level 2
Self-control, -regulation, -configuration, -stabilizing
Maturity level 1
Self-monitoring, -perception, -reflection, -diagnosis, -assessment, (consciousness)

Figure 2.28: Maturity levels of self-x capabilities

capability of adapting automatically and dynamically to environmental changes [159, 160] and ability to automatically install, configure and integrate new software components seamlessly to meet defined business strategies [161].

- **Self-healing** is the ability of a system to automatically detect and diagnose faults, react to disruptions, repair when possible with the objective to maximize availability, survivability, maintainability and reliability of the system [156, 157, 158, 159, 160, 161].

- **Self-optimization** is the ability of a system to measure its current performance against the known optimum [160]. The system always looks for ways to optimize its performance and efficiency in reactive and proactive ways[156, 158, 161] to maximize resource allocation and utilization for satisfying requirements of different users [157, 159].

- **Self-protection** is the ability of a system to establish trust reliably [159], detect, identify and protect itself against various types of attacks to maintain overall system security and integrity [156, 157, 160]. The system uses early warning to anticipate and prevent systemwide failures [158], automatically controls unauthorized access and intrusions, and reports them as soon as they occur [157].

Self-x capabilities are required to reach the higher levels of autonomy in future production systems, and the autonomous nature based on the self-x capabilities are needed to adapt to unforeseen environmental conditions and requirements. Self-x enables autonomy and self-x complexity increases with autonomy level that is directly related to autonomous features of systems[162]. [162] define six levels (from 0 to 5) of autonomy concerning self-x enablement, where level 0 is no autonomy, and level 5 is full autonomy in all areas.

2.4.6 Measurement performance

Measuring performance is critical for any system to understand how well the system is operating, and what can be done to improve it. Many manufacturing devices exist to measure the quality of parts or performance of processes. Assuming data can be collected, measurement performance is a function of accuracy, stability and certainty (see Table 2.7), and it refers to the quality and the consistency of values given by a specific device (being it an instrument, a robot or a whole system).

If uncertainty over the calibration or adjustment of instruments arises, appraising accuracy alone leaves room for errors and misunderstandings.

Failing to assess that a value is stable in different situations and environments, therefore implying that a system cannot be trusted in varied conditions, significantly hinders the utility of a measurement device [163].

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

Table 2.7: Features and functionalities of measurement

Category	Feature	Functionality
Measurement performance	Accuracy	Precision Repeatability Sensitivity
	Stability	Transferability
	Certainty	Calibration

Accuracy

Between the features of Table 2.7, accuracy is typically the first to be considered, but it is also difficult to properly define it. Accuracy encompasses all the qualities that a measurement device–and therefore a measure–must have to be acceptably used in a defined task.

Accuracy is commonly considered the distance between the measured value and the reference value. However, 'precision' is the term that fits this definition better, and it is a just a fragment of what accuracy really defines; therefore, it is the first functionality in the model.

Since ambient conditions might influence the accuracy of a measurement, as a second functionality the model includes 'repeatability', which is the capability to replicate the same measurement—with the same instrument, robot, or system—over a short period of time, and under the same boundary conditions.

The third and last functionality is 'sensitivity', which represents the ratio between a change in measurement value and a change in reference value. If the sensitivity of a measure keeps constant over a range of values, then its transfer function is perfectly linear, and as such it is ideal. However, when calculating the transfer function in practice, it involves—more often than not—some change in sensitivity over time [163]. Sensitivity also includes concerns surrounding the ability to distinguish true positives (TP) and true negatives (TN) from false positives (FP) and false negatives (FN) [164].

In the context of a binary classification, when there's the need to statistically calculate how acceptable a test is, sensitivity is the number of TP correctly identified by the test (i.e. the number of samples that are correctly identified as presenting the desired output) [164].

This would prove useful from an industry point of view, where big data and accuracy of predictions, measurements and instruments will play an ever increasing role [3].

The main purpose of accuracy determination is evident in quality control activities, where a company wants to be sure of realizing products with the highest possible performance level, possibly realizing a defects-free process. The optimization of the process will be based solely on the analysis of the huge amount of measurements that machines will produce. Thus, efficiently and accurately taking the measurements will result in a higher competitiveness on the market [165].

Level one of maturity has no accuracy determination mechanism, but this improves as the autonomy level increases. At higher levels automatic maximization of true positives and true negatives along with minimization of false positives and false negatives is possible without human intervention . At lower maturity levels, the calibration and setting of parameters requires human intervention as the environment changes, but at the highest level five, the system can continuously adapt measurement processes to maintain accuracy despite changing conditions (see Figure 2.29).

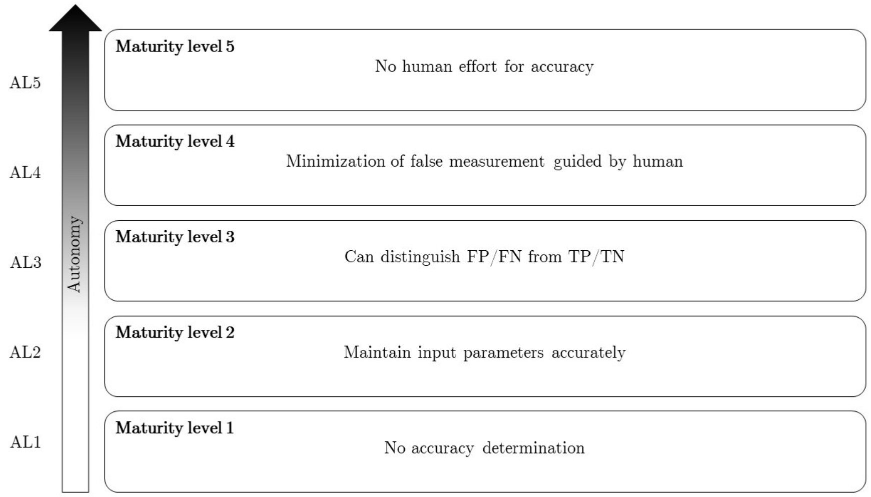

AL5
Maturity level 5
No human effort for accuracy
AL4
Maturity level 4
Minimization of false measurement guided by human
AL3
Autonomy
Maturity level 3
Can distinguish FP/FN from TP/TN
AL2
Maturity level 2
Maintain input parameters accurately
AL1
Maturity level 1
No accuracy determination

Figure 2.29: Maturity levels of accuracy

Stability

Over time, a measurement device could exhibit variations in the quality and reliability of the output value. The alteration could could be due to changing environmental factors [163], and even a total change of scenery—for example when a production autonomous robot is moved to a new production plant (see Figure 2.30).

How can the user be sure that a system that worked with acceptable measurement performances in a previous situation, could keep performing at the same level when changes occur? If a device does that, then it could be considered stable.

A stable autonomous system presents a good level of transferability, which is the ability to transfer performances (for example the accuracy of a well-trained AI model) to other testing domains [166] and that consequently means they are systems capable of making their own decisions and evaluate the challenges that manufacturing proposes, under different and difficult circumstances [167, 168].

Low stability maturity offers no safeguards on the reliability of measurements, or only if the environmental conditions remain constant. At maturity level three systems use artificial intelligence—or some other sort of algorithm—to evaluate its own performances based on historic data of its operations, could face the challenge of being moved to a different environment, where the whole model it has created from previous data-sets is not valid any more.

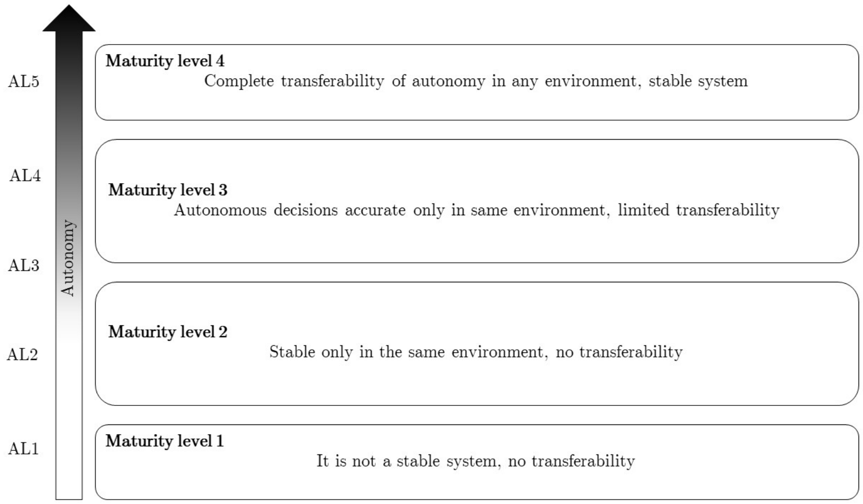

AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 4
Complete transferability of autonomy in any environment, stable system
Maturity level 3
Autonomous decisions accurate only in same environment, limited transferability
Maturity level 2
Stable only in the same environment, no transferability
Maturity level 1
It is not a stable system, no transferability

Figure 2.30: Maturity levels of stability

Certainty

If an operator suspects that the readings of an instrument are inaccurate or deviate from the reference, they could adjust the instrument to correct its sensitivity. This correction could be linear for the simplest devices, but may be non-linear, multi-point for the most complicated systems.

Whenever this adjustment occurs, a calibration of the instrument is needed to verify its accuracy.

All measurements and the calibration of instruments refer to a working primary standard. Therefore, all devices calibrated using that primary standard should be accurate among themselves. The absolute accuracy of every one of the instruments cannot be confirmed unless the calibration uncertainty is defined [163].

Certainty takes into account the ability of production system to validate measurements that are given as an operation input. Production machines are typically capable of taking measurement input for the whole operation cycle; however, they might be unable to distinguish true signals from clutter and noise.

At maturity level one, the system measures values for a given operation cycle. As the maturity level increases, the system might be able to receive measurements for each operation in the cycle, but still without complete certainty.

At a higher level, it takes into account the noise and identifies the factors that cause it. At the highest possible level, the system is capable of dealing with that noise factor by itself: it eliminates the errors to conduct the most precise operation it can. At this stage, no human effort is involved to clean the signal (Figure 2.31).

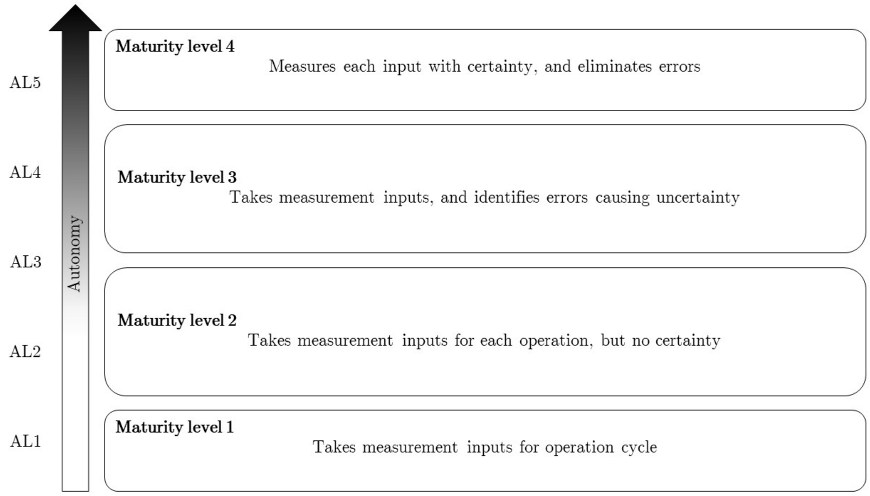

AL5
AL4
AL3
AL2
AL1
Autonomy
Maturity level 4
Measures each input with certainty, and eliminates errors
Maturity level 3
Takes measurement inputs, and identifies errors causing uncertainty
Maturity level 2
Takes measurement inputs for each operation, but no certainty
Maturity level 1
Takes measurement inputs for operation cycle

Figure 2.31: Maturity levels of certainty

2.5 Model Discussion & Conclusions

Section 2.4 stands as a thorough description of the defining model the authors wish to present to the scientific community for autonomy in manufacturing with regards to the requirements and key features, to help break down the complex concept into its constituent parts. It is also intended for companies to be used as an evaluation of their factories' autonomy capabilities, and identification of the missing enablers and requirements. In this sense it is also a road-map from conventional manufacturing systems through to fully autonomous ones.

Enterprises should compare their own achievements–in terms of autonomy–to the definitions given in this paper, and rank themselves, for each of the features/functionalities group, at a specific level of maturity. Then they could understand how far they have come in building an autonomous system, and in which areas there are open requirements.

It is clear that Industry 4.0 has brought digital and intelligent systems to the forefront of manufacturing development, leading to major transformations in manufacturing requirements, capabilities, and operational and management systems. Distributed manufacturing is a key goal that could be achieved through manufacturing digitalisation, and this paper presents an investigation into the ongoing debate around the definition of autonomy in both the scientific and the industrial communities. The contributions of this section are threefold: First, a detailed analysis of the goals of autonomy to help clarify the concept's place among manufacturing systems. Second, the analysis of the features and enabling technologies that can be used to support the identification and implementation requirements of autonomy in the context of manufacturing. Third, the produced five-stage model is designed to give a overall granular definition and guidelines to quantify a particular manufacturing system and pave the way for the gradual transition towards autonomy.

Section 2.2 starts with the high-level definitions of autonomy in literature. The difference between autonomy and automation is also highlighted with the key ability for an autonomous entity to face unanticipated situations and adapt its course of action as they appear. Furthermore, a variety of engineering domains in which autonomy applications were reported in literature is also provided, ranging from unmanned aerial/underwater vehicles to manufacturing areas. Throughout these autonomy applications driven by business objectives in the diverse fields, different autonomous applications will require a different set of features to achieve their objectives, requiring a model to be developed to lay basic for the development of such autonomous systems. This finding opens Section 3 in which different autonomous models and their classification in literature were investigated to understand the current knowledge of autonomy levels along with the key enabling technologies and features.

This leads to the five-level model of autonomy features and requirements in manufacturing systems, from no autonomy to a fully autonomous factory. Low autonomy level manufacturing systems can be highly effective in pre-defined situations, but will depend on human-decision making to react the unanticipated contexts. Conversely, the highest level implies total independent cognitive functioning mechanisms on different features in terms of data, information and knowledge, process, interaction, infrastructure, self-characteristics, and measurement performance.

The system-view of these levels is broken down into the contributing features as indicated by Table 2.2 and Section 2.4 showing the different features associated with maturity levels. As the maturity level of these features and functionalities increase, the overall autonomous level of the manufacturing can also increase. This approach allows the identification of factors which may be holding a system back from achieving its autonomous potential.

Even though there is increasing efforts to perform autonomous transformation in current leading manufacturing industries, such as semiconductor, aerospace, and automotive, as referenced by Figure 2.13, the highest autonomous level has not been achieved. The following work by the authors will be an investigation of the barriers and challenges to increasing the maturity levels of the contributing features of autonomy, as well as an analysis of the business models and considerations which may hold back an otherwise technically possible implementation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

Chapter 3

A Review of Barriers and Challenges to Autonomy for Distributed Intelligent Manufacturing Systems

3.1 Introduction

Each industrial revolution has had its own impact across global aspects for almost every manufacturing and economic sector, as well as societies and matters of environmental sustainability. The first, second, and third industrial revolutions changed industries through making the most of mechanization, electrification and telecommunication - allowing mass production lines to high-level automation [169, 170, 171]. These key technologies have enabled the production of complex products to meet quality and production volume demands of modern societies. However, the changes caused by those industrial revolutions were not all positive and need to be taken into account for the next industrial revolution, especially regarding social acceptance around the automation of jobs, environmental sustainability such as air pollution, social inequality, and need of shaping new roles of humans in increasingly skilled jobs [172, 173]. In addition to the ongoing concerns, the demands of the market today require an increasingly dynamic approach with rapid, diverse changes in manufacturing. The trend of personalisation increases product variants and complexity which is combined with short product life cycles, challenging traditional production systems [13]. Furthermore, technologies change very quickly, posing new challenges in distribution of intersectorial and inter-country connections well as in the structure of labor resources and education. Those changes, unpredictable events, and disturbances cannot be handled efficiently by current conventional hierarchical structures and methods in a productive and cost-effective manner, which stimulates emergence and formation of a new industrial revolution commonly called "Industry 4.0" [171, 11, 174].

Industry 4.0 was announced at the Hanover Fair in 2011 and it is generally regarded as the next industrial revolution based on technologies such as the Internet of Things (IoT) and use of cyber-physical systems embedded in manufacturing factories, enabling the distribution of control and intelligence [175]. These technologies pave the way for the development of factories which are differentiated from the conventional automation by the key ability for intelligent entities to handle changing situations and adapt its course of action at a local level as disruption (or opportunities) appear.

However, as described in the previous section, distributed intelligence is only the first step towards a truly responsive factory. To meet new and unexpected challenges, the system but be *autonomous*. The path to a fully autonomous factory is challenging, which is why we derived the five-level model and definition of autonomy in manufacturing systems, from no autonomy to a fully autonomous factory [176]. Low autonomy level manufacturing systems can be highly effective in pre-defined situations, but will depend on human-decision making to react the unanticipated contexts, even with local distributed intelligence in the assets. Conversely, the highest level implies total independent cognitive mechanisms on different functioning categories associated with contributing features in terms of data, information and knowledge, process, interaction, infrastructure, self-characteristics, and measurement performance. As the maturity level of these features and functionalities increase, the overall autonomous level of the manufacturing can also be enhanced. However, there are barriers and challenges holding companies back from achieving higher maturity levels of these features and functionalities. Even the current leading manufacturing industries, such as semiconductor, aerospace, and automotive, have not achieved the highest autonomous level [177].

Though numerous research efforts have been made to point out the inhibitors to Industry 4.0 [178, 179, 180, 16], a comprehensive study that investigates them at a detailed level while also considering the specific autonomy features is lacking. This is not a trivial task and represents the central contribution of this section. These barriers and challenges imply a wide range of obstacles and issues related to understanding and developing increasingly mature levels of the contributing features of autonomy, as well as the business models and considerations which may hold back an otherwise technically possible implementation. By taking these challenges into account, planning can be made to prepare resources and operations to avoid potential failure as well as maximize the full potential of autonomy in a distributed manufacturing system.

This objective is delivered by providing an analysis of the state of art regarding barriers and challenges: (1) on the path of getting the highest maturity level of each feature and functionality to form fully autonomous factory; (2) on the business models, social and environmental considerations beyond the technical issues; (3) offering future project development perspectives taking those barriers and challenges into account. The following section presents a comprehensive set of technical barriers and challenges structured in their corresponding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

autonomous feature and functionality. Beyond to the technical aspects, Section 3.3 emphasizes on sustainable aspects encapsulating social, economic and environmental barriers and challenges. Section 3.4 ends in its brief conclusion by making the most of the analysis to establish the future project development perspectives in realizing full potential of autonomy in manufacturing context.

3.2 Technical barriers and challenges

When a manufacturer first familiarizes with the concepts of distributed manufacturing intelligence, Cyber-Physical Systems and the Internet of Things (founding concepts of the Industry 4.0 revolution [181]), they might be impressed by the vision of a future where machines will be able to auto-drive themselves and where manufacturing will mostly be driven by machining centers' decisions–with little to no human interaction. That is why managers from important companies have been saying for a few years that they plan to convert production to be autonomous as soon as possible.

However, there are problems linked to a lack of knowledge about the actions that should be undertaken to promote the change, about the technological innovations coming out year by year and the amount of investment required to embrace autonomy [182].

The idea of transforming a company to a totally autonomous manufacturing center is beautiful in theory, but it poses a series of daunting challenges and seemingly insurmountable barriers that are well highlighted by [180]:

- human resources and work circumstances [183, 184];
- shortage of financial resources [185];
- standardization problems [186];
- concerns about cyber-security and data ownership issues [185, 187];
- technological integration [185];
- difficulty of coordination across organizational units [187];
- lack of planning skills and activities [182];
- organizational resistance [182].

While barriers and challenges might often overlap, as a general reference in our concept a technological, economical or resource related issue, that is resolved only with an external action (i.e. a new robot, a conspicuous investment or hiring an expert in the field) should be considered as barriers to the development of autonomy; on the other hand, reluctance and hesitation both by managers and the work force to welcome these new ideas, concerns about security or safety and lack of planning and activities–all elements coming from inside companies that hinders technological advancement–should be regarded as challenges.

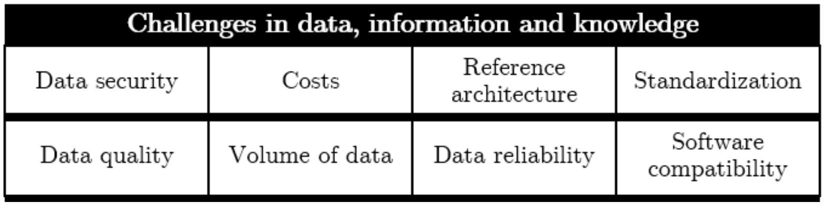
This paper aims to address–at the best of our knowledge–the barriers and challenges that implementing an autonomous factory poses, with regards to the features of the model for autonomy developed by the authors in a previous paper [176].

This section focuses specifically on the technical aspects that hinders the deployment of a fully autonomous manufacturing center, tackling them clustered in the aforementioned features.

3.2.1 Data, information and knowledge

The model developed in [176] identifies 'Data' as one of the dimensions needed to estimate the level of autonomy of a manufacturing system. In general, the data fed to an autonomous system is processed to become information, which is the base for the knowledge, later utilized to make decisions [188]; when this knowledge is fully understood by the system, it becomes wisdom.

Nowadays, companies that understand the power of information as a strategic decision-making tool will be more prepared to face future market competitions [189]; therefore, the aim of this section is to identify the barriers and challenges in data management, based on the features identified in the above-mentioned model, and also summarized in Fig. 3.1.



Challenges in data, information and knowledge			
Data security	Costs	Reference architecture	Standardization
Data quality	Volume of data	Data reliability	Software compatibility

Figure 3.1: Barriers and challenges related to data usage in autonomous systems

Data

Strictly related to data, there are some concerns about the quality, amount, and reliability of the data used to build the knowledge of an autonomous system.

About reliability, the possibility of using corrupt or manipulated data, or that it came from an unreliable source is a big concern [188]. Since this data is the basis for knowledge that the system will build for itself, it is essential that is not contaminated, nor targeted to a specific purpose, as it could generate bias in the system, affecting its decisions [190].

On the other hand, it is possible to assume that a way to ensure data reliability is to analyze significant volumes of it to structure the algorithms or training for the systems; however, this practice does not guarantee that the sample is not biased. Depending on the variables of interest, it is crucial to assure that the sample data is representative enough and carefully selected, to generate valid results [190].

Finally, the data quality is of high importance; the measured variables should have a rational correlation, and be of statistical significance to be suitable to be fed to an autonomous system [190].

The main challenge related to data usage is to develop algorithms and methods to reliably and securely process data so that it will not mislead the system, while also minimizing the need to include a human.

Knowledge management

Due to the increasing value of data in the digital era, another significant concern is its management. Along with collecting the data and having it available to be used, how it is handled, secured, and used is worrisome.

Regarding security, it is necessary to deploy strategies to protect data during its creation, storage, usage, sharing, archive, and destruction. Not only the confidentiality when using sensitive data, but the protection of its integrity is important in a safe autonomous system [191]; this is also related to cybersecurity, which is discussed later in this work, as well as ethical aspects (Section 3.3.1).

Another subject is that the volume of data collected by autonomous systems will require large storage space. Several alternatives should be developed to have quick access to long-time stored data, backups, and recently collected data. Significant efforts on developing storage solutions to allow secure and quick access while maintaining or lowering costs [192] and optimizing energy usage are needed.

Other challenges addressed [193] are related to the processing: the demand for methods and algorithms capable of handling massive flows of data; the laborious task to analyze raw, unstructured data, that is hard to separate values, combine, analyze; and, the difficult and resource-consuming task of cleaning the data to be introduced to a system.

Interoperability

Using the definition given by the the ISO 16100 standard, “interoperability is the ability to share and exchange information using common syntax and semantics to meet an application specific functional relationship across a common interface” [194]. Sharing the information between different domains requires overcoming several barriers, including the ones identified by [195]:

- compatibility of data flows;
- interpretation of data flows in multiple non-standard formats;
- not using existing standard formats to share information.

Information sharing is crucial for achieving autonomous behavior in a system; if the system cannot understand the information coming from different sources, it would not be processed, neither will it be utilized as a system input. Seamless communication between the components of the system requires establishing a reference system architecture and the standards that will govern the information exchange at all levels, which is a task that demands cooperation between different industries and also a substantial investment of time and potentially money [195, 196].

3.2.2 Process

The model developed in previous work places the 'process' category as one of the drivers to ensure autonomic system behavior. This category relates to features that form or influence the physical manufacturing processes of a production environment. End to end physical and logistic synchronisation in production facility along with capability to execute functionality implicitly (functionality is ensured by external inputs but also takes inputs from other preceding processes) of a process directs towards autonomous behavior. This behavior is made consistent by process optimization and reliable execution. Based on the features identified that influence this category, this section describes major challenges and hurdles in the synchronisation of production systems, reaching and optimised state, having reliable execution.

Synchronization

The synchronization aspect in production system autonomy can be categorized into material/logistic synchronization, and physical synchronization as per the work done by [142]. The material/logistic synchronization is concerned with part and material movement from one station to ensure stations are not left idle, and physical synchronisation is concerned with regular controlled part processing at each station.

The major barriers to synchronisation in production system involve failing to achieve ideal working and collaborative conditions, failing to eliminate wasteful movements as well as parts of operation, failing to harness the full contribution capability of the resources and failing to act continuously for self-improvement. The major barriers or hurdles in process synchronisation can be contributed to certain factors which are:

- Lack of sate realisation: The realisation of state of the system by the system itself assists in manipulation of the state for ideal working environment [197]. System setbacks that may result in degraded performance could be controlled by synchronisation of the system at the logistic, cyber and physical levels.

- Lack of end-to-end For achieving complete autonomy end-to-end synchronisation is necessary. The synchronous behaviour could be divided into logistic and physical synchronisation [142]. Logistic synchronisation is relational between component transfer between stations and physical synchronisation is adherent is to all components of the system itself. The system component in this regards should be capable to being made synchronous both in terms of logistic transfer and physical components (the previous components should adapt to the next).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

- Synchronisation in production system can be hurdled by the inability to restrict wasteful activities (a form of noise) [198]. The system should be capable of identifying and eliminating such activities.
- System should be adaptable [199] i.e. acting on best suitable parameters by providing constant comparison through a criteria.

Optimisation

Optimisation in manufacturing involves cost and resource optimisation. Optimisation is about selecting the best possible alternatives among a set that fulfills a specific criterion. The specific criterion in case of manufacturing is all activities that lead to producing of those products which fulfill all required functionalities, performance, characteristics and quality. In a manufacturing environment the optimisation goal must be set at all levels (strategic, tactical and operative). For all levels objectives and constraints must be defined.

At the strategic level optimisation is carried out to support decision making of product/process selection, acquisition of new services and resources. At tactical level optimisation supports production planning, procedure identification and resource management whereas at operative level it deals with decisions of flow management (inventory, scheduling), operation control and quality control.

Cost optimisation leads to a competitive advantage for manufacturers by helping them undercut price relative to competitors, increasing margins, and improving cash-flows. It also leads to reinvestment into the business encouraging potential growth. However, certain challenges and barriers could restrict realizing the potential. One of the contributing factors is poor cost management systems [200]. A good cost management system should enable effective and efficient insight into cash flows at the shortest intervals. These cost management systems should extend all the way from shop floor to top level business and provide a coherent view in terms of financial statements.

Failure to adopt a zero-based budgeting standard [201] may also incur restriction in cost optimization as a zero-based budget enforces a justification on expenditure. This is necessary to continuously challenge the costing paradigm and establishing new cost trajectories. This coupled with performing continuous improvement on budgeting process can lead to effective and robust cost optimization functionality. This continuous improvement in budgeting process is carried out by tightening your budget over time by small increments aided by steps taken to reduce cost. Operating on lower side of cost structure may be beneficial to the organisation and may lead to activities of cost optimization [202].

Cost optimization suffers when fixed and variable costs are managed separately. This could be countered by using an ‘income statement’ approach to the process area level thereby developing a holistic idea of cost. Poor cost allocation may also serve as a barrier for cost optimization as inaccurately associated costs become unmanageable. Costs should always be maintained by cost centers. Cost optimization could also be impacted by poor integration of business goals with costs to optimize the whole business [202].

Resource optimization is essential to ensure production can be carried out to meet varying demands ensuring proper compliance. The major barriers and challenges in resource optimization are usually contributed due to lack of proper information management and processing along with incapability to offer scalability to the varying demand. This in turns restricts the decision-making process and constrains resource optimization.

Resource optimisation deals with a set of processes and methods that match available resources with organisational needs to achieve established goals. The target of resource optimisation is towards minimum utilisation while achieving desired results (under time frame and cost). In a manufacturing systemic approach resource optimisation is linked to concept of constraint and company vision. The company vision is important in case of resource optimisation as it is difficult to define global effectiveness of the resource. The main requirements of resource optimisation are:

- System should have capability to implement and have knowledge on the global goal to be achieved (priority over individual tasks).
- System must have the capability to scale the effectiveness of the task. Match the global goal when performing operations.
- System must identify constraints and provide protection against variation.
- System should have full management capability over process.
- System must have good Statistical Process Control (SPC).

Reliability

A lack of proper standards can serve as a major inhibitor for effective reliability management in manufacturing setting [203]. These standards must provide a set of common definitions and structures for reliability requirements communication, estimation, and testing of parameters and reporting. The lack of a unified framework is another major barrier to Industry 4.0 implementation. Using a framework enables error-free execution by unifying on common understandable terms.

A common unified framework for reliability enables development of common terms which would assist in preventing misunderstandings. The framework might include an outline for data collection, analysis and reporting controlled by an effective communication technology. Novel failure mechanisms can restrict autonomous behaviour due to the novelty of the situation, and that is specially valid in case when such capability to overcome such behaviour is not valid in system.

Reliability may also be effected by changes in environmental conditions along with system use conditions. The system state may also effect the factor as each specific state is influenced by some underlying principle and may perform a function with reliability to a certain extent. For this a clearly outlined method of rubrics/checklist measurement must be present.

3.2.3 Interaction

As the ambition to get a global manufacturing network of autonomous and interconnected systems is getting more and more attraction, it is required to analyse those technological aspects that slow down the improvement of the interactive features of these systems and their deployment in real scenarios. In this section, different barriers and challenges that define the maturity of the enabling technologies are analysed according to different features: i) context-aware technologies, ii) the human interactions, and iii) safety technologies.

Context aware

Context-awareness is one of the key differences between autonomy and automation when used in conventional manufacturing structures. At its highest level of maturity, context-aware systems embed an ability of contextual sensing to aware their own characteristics and constraints and their interaction with entities' preferences and environment characteristics. Such systems accordingly perform contextual resource discovery and adaption by executing autonomous actions to adapt to changes of the sensed environment without explicit intervention [176] [143, 144, 145, 147]. An autonomous manufacturing system is characterized as an inter-connected intelligent system, where the entities are objects which can be a user (humans), place (locations), or thing (machines) whose state regarding a specific aspect or situation is characterized by any information forming the context information. The context itself can also be understood by an active context and a passive context [204, 144]. The active context is needed to identify the current entity and its conditions while the passive context is all other contexts. For instance, if a smart machine intends to provide only the maintenance worker with a service request, the worker´s roles is important and the context represents the active context while the context represents a passive context in which the machine provides all users with a general information and their roles are not matters.

Context-aware systems have been defined, adapted and applied in different domains. One of them is the knowledge context-awareness based systems in which human is supported with the exact needed information flows relative to their skills and expertise, e.g. novice, intermediate and expert workers [143]. Another application published was in the shoe manufacturing process in which the process parameters are autonomously adjusted based on the context to improve error-prone processes and reduce maintenance problems [145]. Many applications have been published [205, 146, 144, 145], showing a tremendous potential of context-awareness systems in the predictability of the work situations, service availability and comprehension.

However, application of context awareness in the manufacturing industry has not yet been sufficiently researched [144]. This is evidenced by the fact that the majority of existing contributions in the field are still in the stage of context sensing without providing a support to adaptation to these new contexts of use. There is also no common architecture providing all context sensing and interpreting components [145, 147]. Several barriers and challenges hold the systems back from getting the highest maturity as full potential of context-aware systems in industrial applications. They can be classified and summarized in data management, network and connectivity infrastructure, and security and privacy management.

Firstly, the issues regarding data management in the context of interconnected processes or sensors are very prominent. They are related to a capability in handling the diversity of data sources and types concurrently, which requires innovative context-aware methods to extract comprehensive contextual information from these diverse data sources [206, 144, 145, 174]. Furthermore, filtering, validation and processing relevant contextual information within the time limits, such as real run-time, are also challenges for the design and development, motivating the rigorous investigation, design, and development of real-time context aware big data processing techniques and the manipulation of contextual information including internal and external entities simultaneously.

In addition to data management, network and connectivity infrastructure must be considered for realizing full potential of context-awareness. The brief and/or intermittent connectivity associated with diverse interaction with entities, which is further challenged by the limited capability of resources, pose a challenge of listing and combining the information of the context states to prioritize actions. Furthermore, another concern lie in the level of sufficient network resources required to be available to such systems at any space and time. Some work is still required to ensure the smooth functioning without requiring intervention [180, 207].

Security and privacy are also prominent issues because context-aware systems are more likely to put critical task and process information over the infrastructure that may be insecure or is not perceived to be secure [206, 207]. This demands a compatible security solution with pervasive mechanisms and a scalable and adaptive policy.

Interaction with humans

In terms of interactions, one of the most challenging concepts of new intelligent manufacturing systems is not only the interconnection or interactions with other systems, but also the introduction of humans into the system. Human roles in autonomous systems are a key enabling factor for autonomous systems [176], and then it is mandatory to analyse and to assess carefully the main barriers and challenges of its introduction.

A goal of future factories is to be better suited for workers with different skills, capabilities and preferences, with greater autonomy and opportunities for self-development. The system will adapt the factory shop-floor to the skills, capabilities and needs of the worker by means of digital assistance systems and technologies. It is defined that two or more actors in a system are in a cooperation if: i) each strives to reach goals while interfering with the others' goals (at least regarding resources or procedures), and ii) they try to manage such interference to make the others' activities easier.

The following sections address the introduction of the human factor dur-

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

ing the design process, and the technological reality when dealing with unpredictable systems, as those in which the human is introduced as a main character.

A Human-centered design

It is common that researchers adopt a techno-centered approach (i.e. giving priority to solve technical issues), favouring the definition and allocation of tasks to automated systems, only taking human operators into consideration at the end of the design process once the control system has been defined, hence giving to the human the role of supervision. This is: to determine the objectives, to tune the parameters, constraints and rules. Even with a "*human-in-the-loop*" approach, in terms of human factors, it can be interpreted more as a way of keeping humans in the "*decisional/responsibility loop*" but not in the "*control loop*": the human operator is excluded from the decisions made by the control system [208]. In addition, this techno-centered approach prioritises automating a maximum number of functions in a pre-defined context and it is assumed that the overestimated ability of the human operator (who behaves perfectly when desired and within suitable response times, as well as reacts perfectly in the face of unexpected situations) will only supervise and handle all the unforeseen situations on the fly. In contrast, in a human-centered approach, the human operator must be involved in the dynamics of the construction of decisions, being aware of the system runtime and an important decisional part of what would be the right rule or parameter to tune. These models are based on following facts [209]:

- The human can be the 'devil', i.e. they may forget, make mistakes, overreact, be absent, or even the root cause of a disaster.
- The human can be the 'hero', i.e. they may save lives through unexpected, innovative behaviours and be an opportunist who can perceive and use unexpected information.
- The human can be the powerless witness, i.e. they may not act despite knowing the right or wrong action.
- The human is legally and socially accountable, i.e. is allocated responsibilities and authorities between humans and machines.

According to the production roles, human-centric functionalities and their technologies are classified in [149]:

- The silent teacher, supporting the learning process of operators.
- Knowledge manager, who handles relevant information flows from different stakeholders.
- Resource integrator, enabling agile and reconfigurable CPPS.
- Caregiver, whose commitment considers the assessment and minimisation of negative impact of certain tasks.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

- Risk manager, whose mission is to create a healthier, safer and productive workplace.
- Flow master, in charge of managing all resources' flows to reduce lead times while improving quality.

Based on these principles, some specifications may be listed to design a more human-centered system [209]:

- The human must be always aware of the situation.
- The levels of automation must be suitable for active human-machine collaboration and adaptive.
- The diversity and repeatability of decisions must be considered.
- The human mental workload must be carefully regulated, as well as the feeling of the need to cooperate and the willingness to do so.

As a solution, new frameworks are proposed that suggest exchanging information along the structure, which is decomposed in to 3 levels [208]: i) operational, ii) tactical, iii) and strategic. In addition, a key proposal is related to the integration of mutual observation to further investigate the limited reliability of the human or either the manufacturing control system. Main streams in this field focus on [210] i) research in automation and task analysis to provide good methods to model the decision-making processes and the boundaries between human and automation control, ii) research on collaboration and symbiosis to provide a useful perspective on the dynamic adaptation and reciprocal support between human workers and technological systems, and iii) labour market to identify jobs that leverage peculiar human capabilities, hardly replaceable by artificial intelligence and automation.

Lastly, these technologies might be deployed to enhance human creative potential as a complement to the robotic and virtual world of the fully automated cyber-physical production system. Enhancing human actuation from the supervisor role ("standby operator") to the maker role.

B Gap expectations between human-centered disruption and unpredictable systems

If we want to be definitive in awarding human operators a central role in industry, much effort must be placed to close the gap between last technology developments and safely control complex and unpredictable systems. Technological enablers of human-centered designs are, among other technological disruptions, supported by cognitive interactions (see Section 3.2.3), advanced visualizations and simulations [149].

A collaborative system must capture the main added-value of work, which includes decision making and problem solving, creative actions and social behaviours, with reference to individual activities rather than jobs [210]. The integration of human factors is enabled through the collection of data about the operator's performance, actions and reactions, aiming at improving the overall factory performance, avoiding bottlenecks. In order to implement this interactive collaboration, a three-step strategy may be followed [210]: i) selection of assembly phases, ii) choice of the level of cognitive automation carrier, and iii) recommendations of cognitive automation content.

Actors' abilities according to each type of task are called *know-how* and *know-how-to-cooperate* [208]. *Know-how* is the actor's ability to control the process by itself, without considering exchanges with other agents. Internal know-how deals with problem solving by actor's skills and capacities, based on its skills for known situations, rules for identified problems and its knowledge when new problems arise and requires analysis. External know-how relates to obtain information from the exterior (process) and the capabilities for actuating over it. *Know-how-to-cooperate* allows interactions with other actors and is also divided in internal (model knowledge of other actors) and external (how to interact with those actors).

Both, augmented cognitive interactions and visual computing technologies may play a key role in empowering operators with digital tools and solutions for improving their decision-making and action-taking processes in specific tasks such as assembly, maintenance, quality control, training, inventory and machine operations [210]. All these technologies enhance operators' ability to perform traditional tasks and to easily learn new ones, not only in isolation but also in connection with other digital and physical systems.

Visual interfaces enable easy communication links which comprise control unit terminals and human-machine interfaces and allow information feedback from and to shop floor [148]. Augmented Reality (AR) devices, monitors on workstations, smartwatches, mobile PCs and tablets are known to be interfaces easily deployed along the whole manufacturing line, to provide instant and intuitive in-situ decision support. Simulations are still a challenge in final assembly operations due to the sequence generation algorithm for the creation of assembly steps, which are computed and then delivered to the aforementioned visual interfaces. Virtual Reality (VR) environments provide a safe and effective environment to carry out human-robot collaboration (HRC) strategies [211].

Since machines and humans share the same physical space, the way in which the autonomous machine (commonly robots) moves and operates, substantially affects the final performance of the whole human-machine team [148]. This represents a trade-off between the human comfort and the low time necessary to the team to perform the assigned tasks. The assessment of ergonomics, emotions and reactions (legibility and predictability, i.e. transparency) during the collaboration has been widely studied, even it still represents a huge challenge in real applications [148, 212, 213].

Safety

This section focuses on the technical barriers and challenges for constructing and implementing safety measures for autonomous systems, that is to implement safe methods and protocols [176]. For that purpose, firstly, general definitions of safe protocols to enhance the implementation of safe and autonomous systems are analysed. The last part of this analysis studies the implementation of moral machines as a solution for safe autonomous systems with enhanced safety protocols (flexible rules).

A Designing safe and autonomous systems: defining protocols

Normally, safety has been analysed as a centralised routine focused on minimizing risks of hierarchical systems by creating stability and control through minimizing uncertainties. As previously seen [176], future autonomous systems are based on decentralized control strategies which are specifically design to manage demands of stability and responsiveness (reactiveness) under external and internal uncertainties (lack of information or ambiguity in available information) by fast and local adaptations and improvisation (the emergence concept). However, that decentralization poses many challenging questions about safety issues [151]. Hence, making traditional safety approaches less suitable. Some research point at the theoretical concept of loose coupling as an approach that meets decentralised requirements for creating mechanisms for these new complex control strategies. The concept of loosely coupling is not easily defined as represents a contradiction which, in general terms, stands for [214] an approach to interconnecting components or elements in a system or network which have either few variables or knowledge in common or in a weak way. That is, elements may affect each other suddenly (rather than constantly), occasionally (rather than constantly), indirectly (rather than directly) or eventually (rather than immediately).

Considering the application of safety measures, one can distinguish between process and personal safety, and between operational and higher-order safety [151]. In process safety, the risks are directly linked to the primary work tasks of the organization (e.g. production and transportation). Potential damages to persons or the environment result from failures in the execution of these primary tasks. Personal safety, potential damages concern the human operators themselves, but they are not necessarily directly linked to the primary work task. Protection against personal safety hazards is mostly a secondary task, sometimes interfering with the primary task. A proposed theoretical framework can be seen in 3.2.

The proposed framework is driven by the motivation of autonomy with the goal of achieving safety behaviours and outcomes to personal and process safety. Uncertainty moderates the whole process mainly creating the need of loose coupling mechanism (concurrent centralization and decentralization) to face dynamic changes. The relation between autonomy and motivation is of important relevance due to its complexity. For example, in the assumption of highly and intrinsically motivating jobs, workers may find a more natural solution for an autonomous decision-making demand under high uncertainty. In contrast, in situations with low uncertainty, keeping workers' participation in reduced operational autonomy decisions (i.e., higher order autonomy) maintains intrinsic motivation, or even require extrinsic motivation (rewards or social recognition). Nonetheless, uncertainty may also affect the relationship between safety behaviour and safety outcomes, the participation in safety becomes more important when more flexibility and adaptation are required, as a response to demands of dynamic and changing situations.

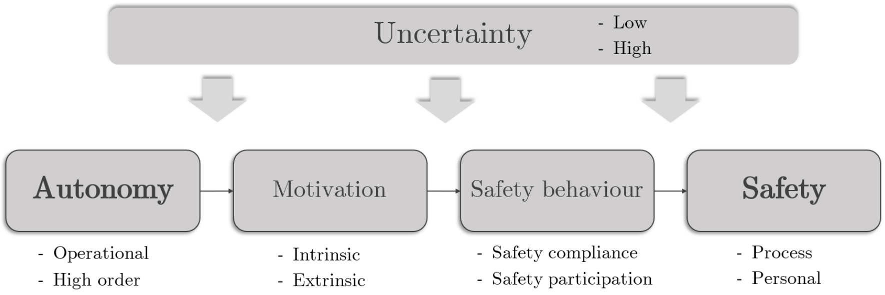


Figure 3.2: Conceptual framework of the relationship between safety and autonomy, based on loose coupling mechanisms and the implications of uncertainty along the whole process. Adapted from [151]

Evaluation methods rank risk and hazards according to their impact (safety) and to the probability of risk (reliability) [215]. For that purpose, it is important to differentiate the terms safety and reliability. For the reliability analysis, readers are referred to Section 3.2.2.

Based on the Quantity Risk Analysis (QRA), a theoretical framework based on Machine Learning is proposed to provide a basis for assurance and regulation of autonomous systems [216]. The most challenging part is that, while traditional systems are only evaluated at the beginning of the systems' conception (prior to deployment) and remains largely valid through life, new autonomous systems require continuous evaluation of safety actions. The framework is divided in four major elements (see Figure 3.3): i) the autonomous system by itself and its operational environment; ii) the design and simulation models, as well as the results of the safety analysis; iii) operational data and results of this data; and iv) the safety/assurance case. Dissonances or distortions between the elements are:

a Between the real world and the world as imaged are:

  - *Assumptions*: Include mismatch between the risk assessment and reality.
  - *Framing (scoping)*: Aspects considered from the real world and which do not.
  - *Model fidelity*: Precision and accuracy of the models used (see Section 3.2.6).
  - *Training data*: Limitations of the dataset used to train autonomous models (see Section 3.2.1).

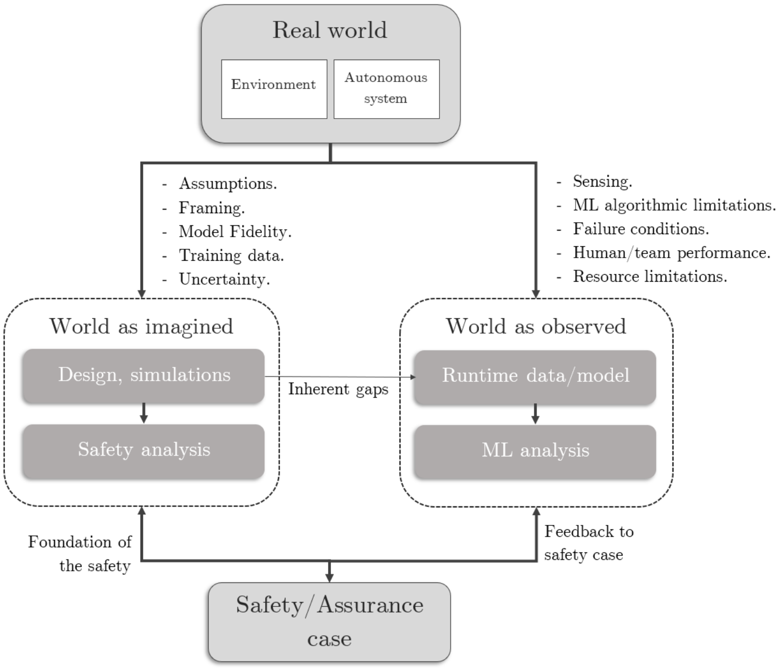


Figure 3.3: General view of the proposed Safety Assurance Framework. The framework continuously observes the system by differentiating between the observer world (collected data from sensors) and the imagined world (safety analysis based on simulations). Adapted from [216]

  - *Uncertainty*: Likely variations in the environment or perturbations unable to be modelled due to the complexity.

b Between the real world and the world as observed are:

  - *Sensing*: Limited performance of the sensors or as a consequence of the environment.
  - *ML algorithmic limitations*: As a characteristic of ML algorithms.
  - *Failure conditions*: Hardware failures affecting the behaviour.
  - *Human/team performance*: Human performance as a result of limitations in cognition and capabilities.
  - *Resource limitations*: Similar to algorithmic limitations but as a result of a lack of resources.

B Autonomous safety: the moral machine

Different mechanisms might be deployed for the construction of safety rules, under the premise of safety first. As an advanced mechanism, flexible rules represent another mechanism for loose coupling as workers are guided and constrained in their actions, while at the same time being called upon to plan and reflect on their actions [151]. In this context, some research works defined the *atomization of actions* paradox, which stands for the motivation to develop a plan - consisting in smaller actions - with the premises of general rules, under the context of a self-regulated goal, following a self-update behaviour in moral machines. Flexible rules propose exogenous limits while encourages exploration and expansion of endogenous capabilities of the moral system.

Under that premise, the development follows many questions: Whose and what morality? (see Section 3.3.1), and How? To answer this last question that relates the technological part, the distinction between artificial moral agents (AMA, as an implicit ethical agent or weak "moral" AI scenario) and artificial autonomous moral agent (AAMA, as an explicit ethical agent or strong "moral" AI scenario) is made by defining the level of "*programmed*" morality [217]. By analysing different moral machine models (Kantian's moral machine [218], Anderson & Anderson's reinterpretation of utilitarism [219], and Howard Mutean's virtue machine [220], pp 121-159), the author concluded that all projects faced the same problems in relating computation to estimation which affect the respective correspondence between epistemic and moral predictability, while building such machines.

3.2.4 Infrastructure

Wide-ranging broadband infrastructure is a must in realizing Industry 4.0. One of the major pitfalls, even in developed nations such as Germany, is a lack of reliable high-speed broadband for SMEs. Industry 4.0 would require every member to be integrated; thus, digital infrastructure is a factor that cannot be ignored.

Connectivity

While data communications have long existed in industrial automation systems, they are built for closed-looped control tasks only. As such, machine and operational data are often trapped within multiple process silos on the factory floor. Connecting these systems via Ethernet can be expensive, cumbersome and may require production shutdowns. In many industrial settings like open-pit mines and oil fields, asymmetric topography and vast geographical areas make trenching wires almost impossible.[221]

The IIoT value chain essentially starts with data collection and choosing the right connectivity solution might ultimately impact the success of your IIoT initiative. There's a plethora of wireless technologies in the market, but not all of them can keep up with the demanding industrial environments. Long-range, deep penetration and high interference immunity of the radio link are key to reliable data connection over large industrial campuses. Also, you'll want to have a unified communications solution to extract data from existing industrial networks and to support a new layer of granular, battery-operated sensor networks for complete operational visibility. In this context, low power consumption and high network scalability are other critical wireless criteria not to overlook.

Cybersecurity

Cybersecurity is a collection of interacting processes intended to protect cyberspace and cyberspace-enabled systems (collectively resources) from intentional actions designed to misalign actual resource property rights from the resource owner perceived property rights.[222]

The exponential growth of the Internet interconnections has led to a significant growth of cyber attack incidents often with disastrous and grievous consequences. Malware is the primary choice of weapon to carry out malicious intents in the cyberspace, either by exploitation into existing vulnerabilities or utilization of unique characteristics of emerging technologies.[223]

The shortcomings that hinder information sharing of cyber threats can be divided into eight categories, identified and described below. Many of these challenges are complex and inter- related, but have been categorized for the purpose of research and analysis.[224]

- **Constitutional / Legal** – Constitutional and legal barriers often prevent firms from sharing information about cybersecurity threats and vulnerabilities. Shortcomings in this area include privacy concerns, as legal protections that mitigate the sharing of personal information(PI)/personally identifiable information(PII) and competitive information are not considered comprehensive or strong enough. Many firms are also impeded from sharing due to concerns about the risk of disclosure of trade secrets, potential legal liabilities and actions that may be taken following the disclosure of cyber threats or attack details, and reputational damage that ensues from these disclosures.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

- **Technological** – Technological barriers include a lack of interoperability or compatibility be- tween the sharing organization and firm. In some cases, organizations do not use a common language, such as Trusted Automated Exchange of Intelligence Information(TAXII) or Structured Threat Information Expression(STIX), to share information about threats and vulnerabilities, but rather create their own sharing language that has not been widely adopted, making sharing increasingly difficult. Without a sharing language that is simple and efficient, fosters uniformity, and creates clarity, sharing is often constrained. Moreover, the complexity of information also acts as an impediment to sharing because the firm does not have the software or hardware capabilities to digest the information being shared, thus making it inoperative.

- **Informational** – Too much shared information and a firm's inability to process this data acts as a significant barrier. This informational barrier makes finding valuable data hidden in a sea of irrelevant noise increasingly difficult to manage. These barriers often include challenges such as unreliable data, the impertinence of information shared, and difficulties of validating data quality..

- **Collaborative** – Collaborative barriers include the challenges of establishing trust between a firm and sharing organization; the process complexity of sharing information; the size of the group information is being shared with; the type of participants receiving the shared information; and a lack of reciprocity from other stakeholders or the problem of free-riders. This barrier category also includes the risk of sharing with rivals/competitors who may use the shared information to enhance their competitive position.

- **Managerial** – Managerial barriers involve challenges around the management of data and relationships from the firm and cyber information sharing organization perspectives. From the firm perspective, these barriers include internal managerial risk aversion and mistrust, often discussed as exposing the firm to "uncontrolled" risks, thus impeding the sharing of information. This managerial barrier could also be called loss aversion, leading to a status quo bias of not doing any more than already is being done. Barriers from the perspective of the sharing organization include the challenges of having no agreement recognizing a single, common centralized authority for establishing trust channels to exchange information between a firm and organization, and a poor management of shared information.

- **Organizational** – Organizational barriers to sharing information include a firm's inability to consume data due to limited resources, and an absence of mechanisms to govern and control the use of sensitive information.

- **Cost** – The costs associated with sharing information about cyber threats and vulnerabilities often serve as a significant barrier. Costs include the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

expenses associated with needing system technologies to share and receive information, which may have a high cost for some firms. Many firms also have limited human resources to process shared data and information, so there are costs with the need to hire more personnel to handle this type of analysis. In addition, barriers in this category include the cost of false positives based on outdated and/or unreliable data provided by a cybersecurity information sharing organization.

3.2.5 Self-X

Self-x capabilities are required to reach the higher levels of autonomy in future production systems, and the autonomous nature based on the self-x capabilities are needed to adapt to unforeseen environmental conditions and requirements. Self-x enables autonomy and self-x complexity increases with autonomy level that is directly related to autonomous features of systems. [191] define six levels (from 0 to 5) of autonomy concerning self-x implementation, where level 0 is no autonomy, and level 5 is full autonomy in all areas. According to the authors, most recent applications of autonomy can be assigned to levels 2 and 3. Non-deterministic behaviour of AI models and the lack of ‘explainability’ of the decisions made by AI algorithms such as Neural Networks hinders the application and implementation of Levels 4 and 5. Hence, AI technologies allow self-x and autonomy, the limitations in current implementations and interpretations of outcomes are technological challenges of enabling self-x capabilities.

Levels of Autonomy with respect to self-x implementation:

- **Level 0** – No autonomy – human beings have full control without any assistance.

- **Level 1** – Assistance with respect to select function – human beings have full responsibility and make all decisions.

- **Level 2** – Partial autonomy in clearly defined areas – human beings have full responsibility and define (some) goals.

- **Level 3** – Delimited autonomy in larger sub-areas – system warns if problems occur, human beings confirm solutions recommended by the system or functions at a fallback level.

- **Level 4** – System functions autonomously and adaptively within defined boundaries – human beings can supervise or intervene in emergency situations.

- **Level 5** – Autonomous functions in all areas including in fluctuating system boundaries – human beings need not be present.

Hierarchy of Cyber-Physical Systems Self-X capabilities.

- **Self-description** (Autonomy level 0): The ability to (formally) describe oneself using a defined language L. With different companies employing different self-descriptions methods, it becomes challenging to integrate devices and types of equipment. To overcome this challenge, there should be an agreement between vendors of industrial machines.

- **Self-monitoring, -perception, -reflection, -diagnosis, -assessment, (consciousness)** (Autonomy level 1): The basis for a more in-depth perception of one's state is to record and know one's state utilizing monitoring, which sets the state of one's system resources (reflection) in relation to the environment or other systems. For this purpose, the ability to diagnose and assess the consequences of an action is part of self-reflection. In this context, the ability to develop consciousness is also called self-reflection of self-reflections in connection with a system's ability to remember. Achieving this level requires constant monitoring of resources and processes. However, the challenge here is 5 V's of big data: volume, velocity, variety, veracity, value. To achieve all these five properties requires technological advances in hardware and software.

- **Self-control, -regulation, -configuration, -stabilizing** (Autonomy level 2): Self-control and self-regulation, as a consequence of a recognized necessary action, is primarily responsible for ensuring that the system can maintain a stable state. The ability to self-configure provides the scope of action in which self-regulation is possible. Self-control is still in the early stages. Trusting machines that autonomously control and regulate themselves is still accepted sceptically. Algorithms must be more reliable, and regulations should change as the methods become more efficient.

- **Self-adaptiveness, -generating, -optimizing, -improvement, -learning, -evolution** (Autonomy level 3): The ability of self-adaptation serves to achieve an optimal operating state (self-optimization) under continually changing conditions and requirements employing self-generated instructions for action. The optimum state can also be improved system-wide or locally. By acquiring new information or capabilities, the system can continuously go through an evolutionary process by which it provides itself with new capabilities. This level of autonomy can be achieved by collecting and learning from experience. However, the main challenge is generating experience in the real-time and physical environment. It is not possible due to technical and safety constraints. Simulation may help make advances, but simulation software should improve as they resemble the physical world as real as possible.

- **Self-protection, -healing, -repair, -servicing** (Autonomy level 4): The self-protection of a system is the ability to arm itself against threats or adverse effects that did not exist during the design. To this end, the system uses self-servicing, self-healing, or self-repair to restore or prevent individual system components or their connection to each other in the event of unexpected disruption or failure. Self-protection can be achieved by collecting possible threats to an individual machine and the whole

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

system. Therefore, some threat monitoring and generalization solutions should be developed to overcome challenges in achieving this autonomy level.

- **Self-organization, -modifying, -modeling, -design, -structuring, -patterning, -assembly** (Autonomy level 5): The self-organization of a system is the ability to alter its internal structure without being influenced by external control elements. This mainly serves to counteract emergence effects, i.e., the occurrence of unforeseen interactions of complex systems. External influences are possible but are considered on a higher level, which reduces the internal complexity of the system to the outside, and only steering influences are allowed. This level of autonomy has a big challenge in integrating data from multiple sources. Artificial Intelligence methods are becoming more robust but for specific data types. However, to achieve this autonomy level, we need fast and robust algorithms to work with various data sources such as images, text, audio, CAD, etc.

Data governance is one of the organizational challenges in enabling self-x capabilities through data-driven technologies which are the basis for autonomy. Data governance includes data quality, data availability, data usability and knowledge, data integrity and sovereignty, data security. Ensuring the high-quality data which is accurate, consistent, and free of ”interference” is a challenging task. Failing to meet the basic quality standards can affect the use and analysis of data-driven algorithms. Different data-driven applications require different formats of data, and making sure that the data is available in requested format is another challenge of data governance. Data usability enables easy search and query. For data usability, data must be structured and documented, which might be a challenging task given the variety of data sources. After data has been gathered and stored, it is important to maintain its essential properties when converted and transferred, i.e., it is challenging to retain data integrity and sovereignty. Given that autonomous systems are connected with each other and with the outside world, data security always one of the important challenges to solve.

3.2.6 Measurement performance

A recent report [225], conducted by the National Physical Laboratory (NPL) in the UK, highlights how the use of autonomous systems is spreading in industry. These systems rely on artificial intelligence (AI) and machine learning (ML) for decisions, for moving and operating; therefore, they need data to be reliable, otherwise they could cause physical and mechanical danger to humans, to machines, to the production systems.

The strategy of NPL for addressing the growth in use of autonomous systems is mostly aimed to assess that transmitted data are of certain quality and integrity standards, be traceable and that they present a defined certainty to be treated as safe and reliable during training and operation. See Section 3.2.1 for extensive discussion about this topic.

In the report the NPL forecast that in the UK:

- by 2025, the global market for telecommunications software will grow to £290 billion;
- Internet of Things (IoT) market will see the number of connected devices rise to 75.4 billion in 2025;
- connected and autonomous vehicles (CAV) could provide an income boost up to £62 billion by 2030, plus a huge leap in job openings and a fall in the number of serious accidents.

Since these aspects of Industry 4.0 are guided by data management, it is imperative that measurement science (metrology) leads the way to build a digital infrastructure supporting digitalisation and that the system stands on trusted digital standards and digital frameworks. This all contributes to enhance the capacity of tracing the chains of data, certificate calibrations and quantify uncertainties [225].

In previous years sensor perception models had already been presented [226, 227]; they aimed to improve the safety of autonomous systems, depending on reliable assessment of the system's context.

One research analysed context awareness by means of two fundamental parameters, safety and precision (it must be noted that the authors of [226] meant "precision" as we intend to define accuracy): safe context meant that the information coming from sensors is not different from reality, while a precise context meant that the information was not "coarser than necessary". It was obvious that an unsafe context might lead to danger or damage to humans and machines, while an imprecise context promotes incorrect operations.

The environment information of this kind of systems always comes from sensors that monitor the area surrounding the machine (i.e. sensors that scan the street for autonomous vehicles), thus this information is an abstraction of reality. The adherence of this abstraction to the real world depends on how accurate the sensor is; it depends on many factors, such as timeliness, trustworthiness, completeness [228].

The evaluation methods for those measures are various [229]: image correlation [230] uses the deviation of the ideal map from the one calculated by the algorithm; path execution [231] computes a distance metric, and it uses both false-positive paths and false-negative paths; therefore, given how accuracy is defined by our previous model, it is better to minimize these false readings.

Vehicle localization used to be driven by reflective markers or electromagnetic guides that helped giving accuracy to vehicle pose [232]; however, given their high cost in terms of money and workforce, the huge number of false positive readings and the fact that those additional infrastructures tend to get dirty or damaged, future developments aim to give vehicles the ability to self-localise without external help [233]. This requires excellent accuracy in scanning the environment and moving the vehicle.

One of the methods to evaluate sensors information for localisation purposes [227] uses the high computational power of modern computers, that was not available before. It has also been tried in a controlled environment with experiments that confirmed accuracy, robustness and reliability. However, standardization has been tried, but without success [234, 235], thus this method–and others–are still not suitable for industrial purposes.

In the same way the field of robotics needs reliable and accurate information from sensors, to build grid-based representation to solve problems like map building, localisation, sensor fusion, path planning and obstacle avoidance [236].

Therefore future improvement will require better navigation and path following algorithms, to raise positional accuracy in order to meet industrial requirements.

3.3 Social, Economic, and Environmental Barriers and Challenges

Innovation represents the most important driver for business growth and competitive advantage, yet manufacturers routinely face challenges around how to successfully manage the process.

In Section 3.2, the authors presented a list of technical barriers and challenges still affecting the deployment of autonomy in manufacturing; however, those are not the only issues that companies have to face when trying to achieve such an improvement [180].

It can be argued that the most difficult barriers and challenges to overcome are social, economical, and environmental. In this regard public exposure, trade unions conflicts, the financial burden that such investments entail, and the environmental impact of any activity must all be taken into account [180, 237].

The following section analyses and tries to better explain barriers and challenges with regards to social, economical and environmental aspects.

3.3.1 Social aspects

As described in a previous work [176], the concept of autonomy is closely linked to decision-making processes which are enabled by massive data collection for analysis, and modeling using artificial intelligence techniques. In this section, such techniques will be studied under their social implications.

The following sections begin by analyzing the acceptance of autonomous systems in the general community and highlighting some commonly known negative aspects. Later, different concepts arise from the analysis of the main barriers and challenges, taken from the concept of autonomous systems, under the perspective of their social impact.

Social acceptance of autonomy

Acceptance in general terms is the process of approving something (tangible or not) newly offered. In the context of social and technological acceptance processes, it also plays a willingness for the use of the technological tool, process or method at issue. This process normally changes over time and is unstable over time. However it not only changes with time, but cultural and societal structures are also important in individual and collective acceptance. Various scientific disciplines (psychology, sociology, marketing and economics) are related to and have mutual bonds with the study of this concept.

The introduction of autonomous systems in society is not always seen in a positive light and it attracts much attention, especially in the contexts of everyday technologies (e.g. autonomous driving systems). Trust in autonomous systems has been analysed from numerous perspectives [238, 239, 240]. Building the credibility of autonomous systems in general society is a very complex issue [241] even inside similar society contexts.

In general terms, knowledge regarding AI and the autonomy concept is still low in society, with a mixture of confused ideas, forecasts of utopias, and science fiction storytelling as many people's only exposure to the concept [240] . It is difficult to correctly define what human intelligence is, hence it is also difficult to apply to machines. Some concepts that are yet to be properly defined theoretically have already been applied in practice. Most people have not yet experienced autonomous systems which encourages a more vague definition [242]. As one of the basic technologies enabling autonomy, AI is normally seen as a "black-box" algorithm, which may result from non-interpretable machine-learning techniques that are very difficult to fully understand. The role of global education and training on general concepts of autonomy and the technologies that enable this feature will normalise the acceptance and will open the way to new applications. In this sense, a great initiative taken by the Helsinki University and the Reaktor Education consisted in globally educate 1% of the European society with basics concepts of AI to demystify its image and empowering new applications[1].

As more autonomous systems have been developed during the last decade and introduced in society (mainly as tests or demonstrators), mass media coverage not only show autonomous systems as a solution to many of current problems, but also show the failures of these systems, with a particular fascination for fatal failures[2] [3], and publicly questioning who has responsibility in the case of a tragedy[4] [5]. However, after further investigations, some reports concluded a wrong use of these systems by the human operators who are actually in change of monitoring these tested systems[6] which actually presents another important challenge to deal.

Some reports sign the way we interact and control autonomous systems is crucial for its real adoption [242]. In the past, research in autonomous systems focused on the technological aspect more than anything. Nonetheless, once the technology is mature enough requires to study social aspects for its introduction

[1] Elements of AI - Link

[2] Wikipedia - Death of Elaine Herzberg - Link

[3] NYT - Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam - Link

[4] BBC - Who is to blame for 'self-driving car' deaths? - Link

[5] Forbes - What Happens When Self-Driving Cars Kill People? - Link

[6] BBC - Uber's self-driving operator charged over fatal crash - Link

in society and to avoid “*disillusionment*”, especially for such applications with high impact in society (e.g. self-driving cars or service robotics). That is the shift to human-centered design approaches and analysis of the socio-technical contexts, focusing on cultural-, type- and milieu-specific differences in acceptance [241].

Another topic that can hinder the social acceptance of autonomous technologies is the riskiness autonomous systems take with its development or the use of autonomous systems for some moral applications or uses. That is the moral use of these technologies for the development of autonomous weapons and its associated dangers[7], forming a negative image of what autonomous systems can actually do for society.

This idea also emerges from the common belief of machines working only under certain discrete states (1/0, black/white, yes/no) and the absence of analytical tools to differentiate wrong conducts from correct ones. Only humans are commonly believed to have moral capabilities or features. When the Three Laws of Robotics were published in Isaac Asimov’s novel[8] it established the initial concept of morality in artificial agents and the roots of a new philosophical current in robotics. More recently, studies analyse the complexity of an artificial morality in artificial agents. In addition, the Turing test[9], and his seminal paper [243]) proposed investigations in this field of artificial intelligence by testing machine’s ability to manifest intelligent behaviours. Some other examples to test this morality problem are the commonly studied trolley problem (based on the dilemma of the double effect problem [244]. More ethical issues of autonomous systems are analysed in the following section.

Ethical issues

Ethical and social impacts of technology need to consider the fundamental intertwining of human and technological domains as both cannot be separated as their interactions are created dynamically in a process referred as co-production [245]. This justified the importance of studying the social ethics on how autonomous systems can be better designed (translation of values), on the participation during the design process, or on understanding the interactions between people and technology. For example, roboethics [246, 247, 248] is human ethics applied to robotics in two levels: adoption of ethical theories, and the robot ethics or machine ethics.

Any application of autonomous system must meet ethical principles. However, from a comprehensive view, many challenges are established by the practical application of ethical principles on autonomous systems. However, many organizations have launched a wide range of initiatives to establish ethical principles for the adoption of socially beneficial systems. This high volume of proposed principles tends to become overwhelming and very confusing. [249] recently reported a unified framework of principles for AI which establishes the

[7] NYTimes - Report Cites Dangers of Autonomous Weapons - Link

[8] Wikipedia – Three Laws of Robotics - Link

[9] Wikipedia – Turing test - Link

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

principles aiming at avoiding overlapping and ambiguity for a practical implementation: i) beneficence; ii) non-maleficence; iii) autonomy; iv) justice; and v) explicability. The work emphasises the concept of explicability as the main enabler for all principles through its division in the epistemological intelligibility concept (how does it work?) and the ethical sense of accountability (who is responsible for the way it works?).

According to [92], in order to have a real impact on society, ethical constraints and aims should shape technology in the design phase in the form of requirements for systems, engineering and software design, as it is in this phase where they actually have an impact. Many works suggest the need for adopting a “fairness-first” approach, instead of evaluating algorithmic bias/fairness as an afterthought [250, 251]. Fairness by design should compel developers to ensure that the very conception and design of AI systems is done in a manner that prioritizes fairness.

Nonetheless, following sections will describe more in detail some of the social practices and principles to face recurrent specific challenges when developing autonomous systems, based on the European Commission’s European Group on Ethics, which established following principles [252], which are: i) human dignity, ii) human autonomy, iii) responsibility, iv) justice, equality and solidarity, v) democracy, vi) rule of law and accountability, vii) security, safety and mental integrity, viii) data protection, and ix) sustainability.

Human dignity

Dignity is understood as the recognition of the inherent human state of being worthy of respect. Human dignity limits the determinations and classifications of persons. It proposes debate about empathy in human-robot interactions (HRI) and whether we can empathise with robots or we should do since they do not have a consciousness in a traditional sense [253].

Human dignity also implies that there have to be limits to the ways in which people can be led to believe that they are interacting with human beings or with algorithms and smart machines. This issue can be found in manufacturing contexts when a human operator directly interacts or is suggested by the system, for example, to meet manufacturing objectives or to follow safety rules (see Section 3.2.3). Beyond the questions of data protection and privacy, we may find an answer to the question if people have the right to know whether they are dealing with a human being or with an AI artefact. Moreover, the question arises whether there should be limits to what AI systems can suggest to a person, based on a construction of the person’s own conception of their identity [252].

Autonomy in control decisions

Considering the capacity of decision-making in autonomous systems, it is important to include security issues which involve: i) external safety for environment and users (see Section 3.2.3), ii) reliability and internal robustness (see

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

Section 3.2.2, and iii) the emotional safety for human-machine interactions and motivation.

Autonomy in control decisions defines how machines honour the human ability to decide when and how to delegate decisions and actions to them. In contrast to the automation of production, it is not appropriate to manage and decide about humans in the way we manage and decide about objects or data, even if this is technically conceivable. Human beings ought to be able to determine which values are served by technology, what is morally relevant and which final goals and conceptions of the good are worthy to be pursued. The principle of Meaningful Human Control (MHC) establishes the constrains the control of autonomous systems and thus who is morally responsible for their decisions (see Section 3.3.1) which, in practice, cannot be left to machines, no matter how powerful they are [252].

Even MHC is mainly studied under the field of automated weapons, it may be also applied in the manufacturing field as there exist potential situations which may affect human integrity. This concept captures three main ideas [92]: a) simple human presence or "being in the loop" is not sufficient condition for being in control of an activity; b) making a substantive contribution to an activity through one's intentional actions might not be a sufficient condition for meaningful control either, which refers to the capacity to respond appropriately under certain circumstances and/or to appreciate the real capabilities of the autonomous system; c) whereas some forms of legal responsibility take simple forms of causal control over events, other forms (typically criminal responsibility) require stricter control conditions of knowledge, intention, capacity and opportunity and, then, the attributions of legal responsibilities of not grounded control conditions turn out to be nor only morally unfair but also difficult to enforce in tribunals (principle of responsibility).

Responsibility

Autonomous systems must be only developed and used in ways that their results and effects align with a plurality of fundamental human values and rights. The philosophical concept of moral dilemmas and common morality apply here, which (very simplified) stand for "some behaviours that are in some sense, generally wrong; and that other behaviours are generally right" in an universal manner [254, 255]. Prior to the development of a technology, it is required to asses the dangers and risks under some developments and researches, potential uses or misuses in future, i.e. the main consequences. One extended example is the moral use of autonomous technologies for the development of robotic autonomous weapons (NYTimes - Report Cites Dangers of Autonomous Weapons). In this case, the moral question emerges: "should the decision to take a human life be relinquished to a machine?" [256]. Possible responses to this question should be analysed at the initial conceptualization or initial steps of their development.

In [92], two conditions are proposed for moral responsibilities assessment as a justification of MHC. The condition of *tracking* refers to the relation between human moral capacities to respond to relevant moral reasons and system actions. In some sense, it refers to the person's total knowledge of the actions taken by the autonomous system and its results (or truth-tracking account of knowledge). In contrast, the *tracing* condition tries to capture the basic intuition that a human agent may be responsible for an outcome even if he does not satisfy the condition for responsibility at the time of the action.

Other extended examples arise with the endowment of ethical mechanisms to artificial cognitive agents as a response of some issues likely to arise (i.e. the implementation of flexible safety rules, see [176]) in the relationship between these autonomous systems and human beings [257]. Machine ethics is a field "*concerned with ensuring the behaviour of machines toward human users, and other machines as well, (and what) is ethically* " [219]. As previously discussed, the definition of artificial moral agents (AMA) and artificial autonomous moral agent (AAMA) is made [217]. Even those machines are technologically realizable, the role of the Moral Lag Problem raises some additional concerns: if humans cannot be as moral as they should or could be, how can machines be moral? Even humans can make up diverse moral claims in contrast to machines that are left with a static account of morality as they would only be able to simulate, rather than to make judgments (lack of pluralism of human moral judgements). Nevertheless, some authors defend the thesis that the rise of AI and robots will "suppress our moral agency and increase the expression of moral patience" as a result of the tendency to suppress the expression of our capacity for moral agency and accentuate the expression of moral patiency [258].

Furthermore, cognitive tools and actions no longer need to be programmed by humans. Some examples of these "self-programming" algorithms are Google Brain (Google Brain Team), which builds AI better and faster than humans; or how AlphaZero develops itself from scratch in 24 hours how to manage some games' rules beating the best players (AlphaZero: Shedding new light on chess, shogi, and Go).

In addition, is of the human responsibility to ensure basic preconditions for life, preservation of a good environment for future generations, ensuring the priority of environmental protection and sustainability. This topic is deeply covered on Section 3.3.3.

Traceability, liability and accountability

Besides moral responsibility, a big concern is the legal responsibility: who should be found accountable for the actions of an autonomous system? who is to compensate the damages for such actions? In the case there is an operator behind the machine or system, is the manufacturer legally responsible? or is it the operator?

Statements like [252, 192] claim that because of the nature of the artificial systems, some aspects of human agency could not be applied to them. Despite that, the concept of liability, blame, and accountability should be extended to include these emerging systems and to enable the application of insurability as well.

To avoid addressing these matters unevenly, it is urgent to create a global legal framework to regulate autonomous systems, protect human rights, and implement effective mitigation systems.

Justice, equity and solidarity

This principle defines global justice, equality in access, and fair distribution of the benefits and advantages that autonomous systems bring. Additionally, not only is important the advantages when an autonomous systems is deployed, but also the access to education as previously exposed is an important factor for its acceptance in society (see 3.3.1). Vigilance is required with respect to the detailed data on individuals which could put pressure on the idea of solidarity, and which may undermine social cohesion and give rise to radical individualism [252].

As one of the main key enabling factors of autonomous systems [176], trained models based on data and AI artefacts, like most human creations, tend to reflect the goals, knowledge, and experience of their creators. They also draw from the strength and weakness of the data that is used to train them and how it is trained [250]. Statistical models, including those created with machine learning techniques, can reproduce biases taken from the historical data used to train them. The use of this techniques by institutions (enterprises or managing levels at manufacturing companies) to automate decisions that affect people's (suppliers, clients or operators) rights and life opportunities then reflect these biases.

As per the bias definition (as a deviation from standard) usually has a negative connotation. So that, it is understood as something to be avoided because of its problematic nature. However, the statistical biases might be neutral but a sign of a different concept, moral biases. These observations are relatively uncontroversial on their own, but they already present a problem for the notion of "algorithmic bias" [94]: there are multiple, different categories of bias that are often treated as equally problematic or significant, even though not all forms of bias are on a par. Researchers have begun developing new techniques to help detect and address these biases as consequence of ML trained datasets lacking in diversity and inclusivity of minorities, as it is one of the most important challenges that AI techniques faces.

On the other hand, many machine learning fairness practitioners rely on awareness of sensitive attributes i.e. access to labelled data about people's race, ethnicity, sex, or similar demographic characteristics to test the efficacy of debiasing techniques or to directly implement fairness interventions [259]. The use of computing solutions can help practitioners and analysts to measure long-standing social problems and to diagnose how they manifest [260].

It is vital to be aware of these risks and minimize potential biases, for example, when deciding [261]: i) which ML technique or procedure to use, and ii) what data set they want to use. Biases not only may result in unfair results regarding sex, gender, ethnicity or skin color; but also, regarding age or disabilities.

Some sources of algorithmic bias are [94]:

i Training data bias, through deviations in the training or input data provided.

ii Algorithmic focus bias, through differential usage of information in the input or training data, e.g. the use of morally irrelevant categories to make morally relevant judgements.

iii Algorithmic processing bias, e.g. by using a statistically biased estimator in the algorithm.

iv Transfer context bias, when the algorithm or resulting model is employed outside of their particular context of operation.

v Interpretation bias, or misinterpretation of the algorithm's outputs or functioning by the user, or by the broader systems.

So that, in general terms, efforts to mitigate biases must consider following steps [94, 250] :

i Creating cross-disciplinary teams of data scientists and social scientists.

ii Identify problematic bias, by assessing whether a given bias is problematic when all things are considered.

iii Intervention on problematic bias, understanding the relationship between the autonomous system and the ethical and legal norms for the relevant contexts. In case no defined norms exist, build fairness measures into the assessment metrics of the program.

iv Ensuring that there is a critical mass of training samples so as to meet fairness measures.

v Adopting debiasing techniques.

Considering last points, race and gender biases are further analysed according to the accessed literature.

- **Race.**

  Apart from the aforementioned general considerations, current methodologies fail to adequately account for the socially constructed nature of race and, instead, they adopt a conceptualization of race as a fixed attribute [262] . The concept of fairness, not only is situational, evolving, and contested from a number of philosophical and legal traditions, but also can only be understood in reference to the different social groups that constitute the organization of society. Consequently, the vast majority of algorithmic fairness are specified with reference to these social groups, often requiring a formal encoding of the groups into the dataset or algorithms, defined as "*protected classes*", and whose construction process is fraught: the construction of categories has a long history of political struggle and legal argumentation.

General tendencies are identified when using race in algorithmic fairness research: i) simplistic conceptualization of race as a single dimensional variable that could delete the complexity of racial categories; and ii) the little attention put on the process conceptualising and operationalising for mitigating differences. Main implications for using race in algorithmic fairness [262] require the multidimensionality of race categories conceptualization and their optimization, to perform disaggregated analysis to understand and identify markers of social inequality, to understand the frame limitations and, finally, to disseminate these decisions together with the main results.

- **Gender.**

  Some works expose that this problem appears even at early stages of some research fields, AI research, for example. Some dimensions of the gender bias that can end up strengthening society's biases are [250]:

  i *Under-representation of women in engineering and research fields.* For example, when the Institute of Electrical and Electronics Engineers (IEEE) instituted a Hall of Fame to acknowledge the leading contributors to AI, no one of the ten persons on the list was a woman

  ii *Perception and stereotypes of the real world*, which are built from the data coming from the world and is used to train algorithms. In this sense, a solution for stereotypes is the neutralization of language suggested to enhance fairer outcomes in natural language processing.

  iii *Disregard of trans and non-binary people.* If gender bias against women is still a problem, what can be expected in the case of other genders, which are not even considered by tech companies? Even worst is to be aware that these companies do not control how gender identity is used in their algorithms [263].

  Proposed solutions tend to improve the representation of women in AI research (as researchers and beneficiaries), from the perspective of "*fairness by design*".

Data protection and privacy

Includes the right to protect personal information and the right for privacy. With regard to the implications of autonomous systems on private life and privacy, two new rights are under debate: the right to meaningful human contact and the right to not be profiled, measured, analysed, coached or nudged [252]. Specifically in the manufacturing field, these issues might appear when monitoring operators behaviour during production tasks for safety purposes or for the establishment of individual labour ratios or performance indices.

In addition to the challenge of facing algorithm biases (see Section 3.3.1), ethical problems in decisions surrounding the data collection and annotation process still remains an overlooked part of the machine learning pipeline [264]. Archives, libraries and datasets are fields dedicated to human data collection for posterity which have grappled with questions of ethics, power, privacy and, specially, consent. Some works summary of critical needs for fair practices among industry ML practitioners and identify the lack of industrial standard for "*fairness-aware data collection*". The lack of any systematic process for generating datasets has spurred researchers to call it the "wild west" [265].

In a collaborative manufacturing ecosystem, crowd-sourcing [264] is a staple approach in collecting human labels for datasets aimed to reduce costs and speed up data collection by outsourcing to human participants are constructed. Crowd-sourcing mechanism provide a set of fixed labels for participants to choose from, which imply adversarial relationships between the researchers and the participants and face criticism for imposing labels onto individuals. So that, the aim is to widen the channel of user input in data collection with their consent, to democratize the collection process and give agency to minority groups to represent themselves. Data power or ownership has been under discussion as well [264], resulting in the development of consortia models, which are set up by institutional frameworks and services to share resources. The main advantage is the ability to gain economies of scale. However, has been under criticism for creating bureaucracy, unnecessary committees, delay and new forms of power imbalances which hamper their realistic implementation.

Techniques to capture users' attention or induce addictive behaviours are now recognised to be at the core of the business models for the most powerful technology companies (The guardian - 'Our minds can be hijacked': the tech insiders who fear a smartphone dystopia), which are increasingly relying on AI/ML systems to collect and extract economic value from any type of data.

Standards become more difficult to implement as objectives move farther from final market transactions and ethical practices in data collection are often overlooked from the distance. The archival codes of ethics list the core values of archivist to be to promote access, ensure accountability and transparency. In practice, archivists are not positively rewarded but only under pressure of code of conduct in their work organizations. In addition, a cross-institutional organization may help ensuring that ethical principles withstand profit-driven motives.

In response, [264] propose to take actions at macro and micro levels:

- *Macro* (community, private institutions, policy makers and government organizations):

  i Congregate and develop data consortia.

  ii Establish professional organizations to work by membership to improve introduction to ethical guidelines.

  iii Support community archives.

  iv Develop a subfield dedicated to the data collection and annotation process.

- *Micro* (individual researchers, practitioners and administrators):

  i Define and modify mission statements.

  ii Hire full-time staff on data collection tied to professional standards.

  iii Work towards public datasets.

  iv Adopt documentation standards keep it rigorously.

  v Develop substantial development policies with domain expertise and nuance of data source.

  vi Make informed committee decisions on discretionary data.

Democracy in policies

To solve most of the aforementioned barriers and challenges, democratic debate and public engagement (in a global spirit of cooperation) should guide key decisions on the regulations of AI developments and applications. Through education and access to information, everyone could understand risks and opportunities and become empowered to participate in decision-making processes for the society of the future [252].

For example, the European Commission built ithe European Group on Ethics in Science and New Technologies (EGE)[10] in 1991, aiming at providing high quality, independent advice on ethical aspects of science and new technologies in relation to EU legislation and policies. In 2015, the United Nations Interregional Crime and Justice Research Institute (UNICRI)[11] launched its programme on AI and Robotics to cope with the legal, ethical and societal concerns and challenges of artificial intelligence and autonomous systems.

3.3.2 Economic aspects

The basis of existence of any organisation is anchored in economic profitability [266]. This means ensuring good monetary returns, liquidity and profitability on production. This in turn should outweigh the cost of investment. Therefore, companies before investing ensure long term economic sustenance and create business cases to justify capital expenditure [267].

A major economic concern with the implementation of Industry 4.0 aspects in manufacturing environments is that currently very few companies have mature digitalization capability along their value chains (both horizontal and vertical) [196]. The digitalization capability enables business services and value chain outsourcing. The main drivers for adopting technology in manufacturing industry is a desire to enhance productivity, production flexibility, and increase output. [268] cites the main motivation of Industry 4.0 to be the development of new services and models for leveraging competitive advantage. The promotion of customer loyalty to the product, quality improvement and capability to predict faults and risk in the product are also essential to the adoption of Industry

[10] European Group on Ethics in Science and New Technologies (EGE) - Link

[11] UNICRI - Centre for Artificial Intelligence and Robotics - Link

4.0. However, in spite of the explored advantages the underlying requirement of infrastructure complexity, huge investment costs and data privacy hinder the implementation of the principle.

[180] explores the main challenges to implementation, listing the lack of available resources, skill-gap, infrastructure unavailability, incompatible processes, variation in demand, data security, compatible solutions and legacy technology as the main barriers. Economic challenges for manufacturing industries are concerned with the economic viability or justification of the required investment for the implementation. In some cases the investment costs is assumed to be on higher end even than the company growth per term. This makes it difficult to provide a solid rationale for such investment. Also, the promise of resulting sales growth in relation to is not always strongly supported by data.

[180] factors in financial resources and profitability to establish the driving forces and barriers in terms of economic understanding. The cost reduction following decrease in human resource cost, inventory management and operation cost is barred by lack of financial ability, limited knowledge of return and profitability as a result of investment and shortcoming of procurement process. Mostly the procurement process in industry revolves around quoting and tendering process which could be time consuming and non-linear. Setting up a standard approach to this process may result in economic backlog.

The research done by [269] studied the implications of Industry 4.0 realisation in current manufacturing environments. The research highlighted the increase in efficiency and productivity witnessed due to smart manufacturing. They found a lack of investment plans and practical guidelines for Industry 4.0 implementation as the major hurdles. These need to be developed at regional and national levels to support effective transition of current manufacturing infrastructure to a smart manufacturing environment. Along with this with stakeholder engagement can play a critical role in development of these guidelines [270]. The marginalisation of regional space is important for integration at that level and then from that into a central economic structure.

The other economic implications of Industry 4.0 include work environments, skill development, growth in potential, sustainability, along with digitalisation and application of intelligence in manufacturing.

- With respect to work environments, automation and robotics may replace a lot of jobs performed by labour leading to economic shift. A significant skill change paradigm will enable new employment opportunities bringing a net employment increase [271].
- Skill requirement will also import economic outlook. At the initial level it will require transition however this transition promises long term incentives[12]). There will be a major skill requirement in networking, IT and IoT. The data requirements of industry 4.0 would assert more effective cyber security and data architectures within the system. This will directly effort the economic outlook of the company.

[12] https://www.controlglobal.com/blogs/guest-blogs/must-have-skills-for-industry-4-0/

- With new data collection, manipulation and decision making capability new growth potential may be witnesses as products and parts become smarter i.e. capable of making decisions about themselves.[205]
- Digital twins [272] can simulate conditions of the part or product to help understand the maintenance of the product to match long term economic goals. However, these require long term investment, communication capability and computing requirements.

3.3.3 Environmental aspects

The environmental aspects of industrial production were once considered as a source of unnecessary expenses, but many steps forward have been made in the last thirty years. Companies have been driven to think more about how to handle resources, emissions, and waste. Moreover, they have started to realize that being environmentally friendly could even be a source of profit [273].

Being known to the public as someone who cares about our planet's survival gives people one more reason to become customers, especially now that awareness about environmental problems is high. On the other hand limiting the cost of resource wasting and of useless energy consumption could constitute a direct source of income.

Therefore one recent focus has been waste reduction and decrease of material usage, along with water and energy use control [274].

However, now that industry has reached a reasonable level of environmental conscience, the new challenges of Industry 4.0 have risen and become the focus of future development.

Perhaps the most important challenge is to contain and understand how to reduce the speed of climate change [275], along with trying to efficiently manage natural resources and energy sources. Companies have the biggest role in avoiding being careless with the environment. They need to develop a sustainable mindset, aiming at addressing climate breakdown and resource scarcity, and creatively look for sustainable solutions to production problems [276, 275].

Industry 4.0 is driven by data management and data gathering [181], and as a consequence other critical environmental initiatives are focused on the potential that this huge amount of information has over green tasks.

According to [277], potential improvements comes from better data quality and less discretion for management over what is measured and how it is reported, since data is always available in real-time. Moreover it means that previously unknown information could be gathered and that their transfer should be smoother and quicker.

Cyber systems, dominated by sensors' presence, could record material and energy consumption by the minute, and find patterns that highlight useless waste on which companies could act.

Again, [277] writes that potential activities in the future could be:

- establish which sectors will be most sensitive to environmental activities impact;

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 814078

- examine the size of companies most affected by Industry 4.0;
- look at the impact on roles and communications between managers;
- what impact will managers have on environmental issues accounting;
- highlight the benefits that environmental accounting have and promote this activities.

Therefore the focus should shift towards making it easier to tackle environmental challenges by collecting and analysing huge amounts of data, and to show what impact environment related activities could have on accounting in companies, so that they would be more interested in investing: this would create a virtuous circle that should bring long term benefits as far as emissions and waste are concerned.

3.4 Barriers and Challenges Discussion & Conclusions

It is now clear that the way from a conventional manufacturing system, through to distributed intelligence implementation, and into a fully autonomous factory is fraught with several barriers and challenges.

Since the first Industrial Revolution almost 250 years ago, manufacturing businesses have faced almost constant disruption – from geo-political turmoil and economic decline, to technological advancements and fast-changing customer needs [278]. Any barrier in the field of digital transformation leads to a slow down or complete termination of the digital change in enterprises.

In order to achieve a more effective production, these barriers and challenges should be faced and overcome. As it is often the case, bigger multinational ventures should be at the head of the evolution of manufacturing system towards maturity; thanks to their superior economical power, they should consider the superior driving forces that make this development worth the investment, and also draw a path for smaller companies to follow.

As seen along the work and apart from the benefits autonomous systems might bring to the manufacturing ecosystem, technological developments have social, economical and environmental implications that must be considered to meet responsible research and innovation actions, supported by the European Commission[13]. Especially, all those changes with high impact on society as it does the new I4.0 paradigm. Previous sections analysed some actions to face all challenges and barriers. However, the general discussion presents the strong necessity of building global and democratic policies to support and to encourage the implementation of all technological developments. Only with a responsible research and a common implication, society will succeed in its most ambitious propositions.

[13] European Commission - H2020 - Responsible research & innovation - Link

Bibliography

[1] A.W. Colombo et al. “An agent-based intelligent control platform for industrial holonic manufacturing systems”. In: *IEEE Transactions on Industrial Electronics* 53.1 (2006), pp. 322–337.

[2] James P Womack et al. *The machine that changed the world: The story of lean production–Toyota’s secret weapon in the global car wars that is now revolutionizing world industry.* Simon and Schuster, 2007.

[3] Henning Kagermann et al. “Securing the future of German manufacturing industry: Recommendations for implementing the strategic initiative INDUSTRIE 4.0”. In: *Final report of the Industrie* 4.0 (2013).

[4] Hugh Boyes et al. “The industrial internet of things (IIoT): An analysis framework”. In: *Computers in Industry* 101.December 2017 (2018), pp. 1–12.

[5] Luis Ribeiro et al. “Transitioning from Standard Automation Solutions to Cyber-Physical Production Systems: An Assessment of Critical Conceptual and Technical Challenges”. In: *IEEE Systems Journal* 12.4 (Dec. 2018), pp. 3816–3827.

[6] Feng Lin et al. “Decentralized control and coordination of discrete-event systems with partial observation”. In: *IEEE Transactions on automatic control* 35.12 (1990), pp. 1330–1337.

[7] Yoram Koren et al. “Reconfigurable manufacturing systems”. In: *CIRP annals* 48.2 (1999), pp. 527–540.

[8] Claudine Badue et al. “Self-driving cars: A survey”. In: *Expert Systems with Applications* (2020), p. 113816.

[9] Katja Windt et al. “Autonomy in production logistics: Identification, characterisation and application”. In: *Robotics and Computer-Integrated Manufacturing* 24.4 (Aug. 2008), pp. 572–578.

[10] Antonio Maffei et al. *From flexibility to evolvability: Ways to achieve self-reconfigurability and fullautonomy.* Vol. 42. 16. IFAC, 2009, pp. 74–79.

[11] B. Scholz-Reiter et al. “Autonomous Processes in Assembly Systems”. In: *CIRP Annals - Manufacturing Technology* 56.2 (2007), pp. 712–729.

[12] Eckhard Hohwieler et al. *Intelligent production systems in the era of industrie 4.0-Changing mindsets and business models Holistic optimization of sculptured surface manufacturing View project Fraunhofer Innovation Cluster Maintenance, Repair and Overhaul View project internet-of.* Tech. rep. 2. 2017.

[13] Olivier Cardin et al. "Evolution of holonic control architectures towards Industry 4.0: A short overview". In: *IFAC-PapersOnLine* 51.11 (2018), pp. 1243–1248.

[14] Thomas Gamer et al. "The autonomous industrial plant - Future of process engineering, operations and maintenance". In: *IFAC-PapersOnLine.* Vol. 52. 1. Elsevier B.V., Jan. 2019, pp. 454–460.

[15] Hong-Seok Park et al. "Autonomy for Smart Manufacturing". In: *Journal of the Korean Society for Precision Engineering* 31.4 (2014), pp. 287–295.

[16] Roland Rosen et al. "About the importance of autonomy and digital twins for the future of manufacturing". In: *IFAC-PapersOnLine* 28.3 (2015), pp. 567–572.

[17] Mica R Endsley. "From Here to Autonomy." In: *Human factors* 59.1 (Feb. 2017), pp. 5–27.

[18] Hanna Theuer et al. "The impact of autonomy on lean manufacturing systems". In: *Lecture Notes in Mechanical Engineering.* Vol. 7. Springer Heidelberg, 2013, pp. 1413–1423.

[19] Luc Steels. "When are robots intelligent autonomous agents?" In: *Robotics and Autonomous Systems* 15.1-2 (July 1995), pp. 3–9.

[20] Alvaro Moreno et al. "The autonomy of biological individuals and artificial models". In: *BioSystems* 91.2 (Feb. 2008), pp. 309–319.

[21] Sarah Buss et al. *Personal Autonomy.* Ed. by Edward N. Zalta. 2018.

[22] Ewart J. de Visser et al. "From 'automation' to 'autonomy': the importance of trust repair in human–machine interaction". In: *Ergonomics* 61.10 (Oct. 2018), pp. 1409–1427.

[23] Jenay M Beer et al. "Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction". In: *Journal of Human-Robot Interaction* 3.2 (June 2014), p. 74.

[24] Riccardo Gervasi et al. "A conceptual framework to evaluate human-robot collaboration". In: *International Journal of Advanced Manufacturing Technology* 108.3 (2020), pp. 841–865.

[25] Angelika Styr et al. "Description Model for the Assessment of Autonomous Production Stages". In: *Procedia CIRP* 93 (2020), pp. 353–358.

[26] Pieter Abbeel et al. "Toward a Science of Autonomy for Physical Systems: Paths". In: (2016).

[27] Panos J Antsaklis et al. “Control and Machine Intelligence for System Autonomy”. In: *Journal of Intelligent & Robotic Systems* 91 (2018), pp. 23–34.

[28] Jerzy W. Rozenblit et al. “Towards design and control of high autonomy manufacturing systems”. In: *Proceedings of the 3rd Annual Conference of AI, Simulation and Planning in High Autonomy Systems: Integrating Perception, Planning and Action.* Institute of Electrical and Electronics Engineers Inc., 1992, pp. 38–45.

[29] Norbert Gronau. “Determinants of an Appropriate Degree of Autonomy in a Cyber-physical Production System”. In: *Procedia CIRP* 52 (2016), pp. 1–5.

[30] Michael Fisher et al. “Verifiable Self-Certifying Autonomous Systems”. In: *Proceedings - 29th IEEE International Symposium on Software Reliability Engineering Workshops, ISSREW 2018* (2018), pp. 341–348.

[31] Andrew Williams. “Defining Autonomy in Systems: Challenges and Solutions”. In: *Autonomous Systems : Issues for Defence Policymakers.* Ed. by Andrew P. Williams et al. OCTOBER. NATO OTAN Allied Commander Transformation, 2015. Chap. 2.

[32] P. A. Hancock. “Imposing limits on autonomous systems”. In: *Ergonomics* 60.2 (Feb. 2017), pp. 284–291.

[33] Tariq Samad et al. “Autonomy in automation: trends, technologies, tools”. In: *Computer Aided Chemical Engineering* 9.C (Jan. 2001), pp. 1–13.

[34] Bernard P. Zeigler. “High autonomy systems: Concepts and models”. In: *Proceedings. AI, Simulation and Planning in High Autonomy Systems.* Publ by IEEE, 1990, pp. 2–7.

[35] Nikolaos P. Ventikos et al. “A systems-based application for autonomous vessels safety: Hazard identification as a function of increasing autonomy levels”. In: *Safety Science* 131 (Nov. 2020).

[36] Daniel J. Fagnant et al. “Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations”. In: *Transportation Research Part A: Policy and Practice* 77 (July 2015), pp. 167–181.

[37] Luciano Floridi et al. *On the morality of artificial agents.* Aug. 2004.

[38] Somaiyeh MahmoudZadeh et al. *Autonomy and Unmanned Vehicles.* Cognitive Science and Technology. Singapore: Springer Singapore, 2019, p. 116.

[39] H.R. Widditsch. “SPURV-The First Decade”. In: *APL-UW 7215, Applied Physics Laboratory University of Washington* (1973).

[40] Sarah Witman. “Detecting Gas Leaks with Autonomous Underwater Vehicles”. In: *Eos* (2017).

[41] Hiroshi Iwakami et al. “Approaching whales by autonomous underwater vehicle”. In: *Marine Technology Society Journal* (2002).

[42] E. An et al. “Coastal oceanography using a small AUV”. In: *Journal of Atmospheric and Oceanic Technology* (2001).

[43] Stefan B. Williams et al. “AUV Benthic Habitat Mapping in South Eastern Tasmania”. In: *Springer Tracts in Advanced Robotics* (2010).

[44] *Houston-based company intuitive machines.*

[45] Paolo Gunetti et al. “Simulation of a Soar-Based Autonomous Mission Management System for Unmanned Aircraft”. In: *Journal of Aerospace Information Systems* (2013).

[46] Vladimir Djapic et al. “Using collaborative autonomous vehicles in mine countermeasures”. In: *OCEANS’10 IEEE Sydney, OCEANSSYD 2010.* 2010.

[47] Somaiyeh Mahmoud.Zadeh. “Autonomous Reactive Mission Scheduling and Task-Path Planning Architecture for Autonomous Underwater Vehicle”. PhD thesis. Flinders University of South Australia, 2016.

[48] Md Jahidul Islam et al. *Person-following by autonomous robots: A categorical overview.* Vol. 38. 14. 2019, pp. 1581–1618.

[49] Trung Dung Ngo et al. “Multiagent robotics: Toward energy autonomy”. In: *Artificial Life and Robotics* 12.1-2 (2008), pp. 47–52.

[50] Georgios Andreadis et al. “Classification and review of multi-agents systems in the manufacturing section”. In: *Procedia Engineering* 69 (2014), pp. 282–290.

[51] Luka Peternel et al. “Teaching robots to cooperate with humans in dynamic manipulation tasks based on multi-modal human-in-the-loop approach”. In: *Autonomous Robots* 36.1-2 (2014), pp. 123–136.

[52] Xianjing Yang et al. “Image recognition of a bolt in power distribution line maintenance tasks for an autonomous robot”. In: *2013 International Symposium on Micro-NanoMechatronics and Human Science, MHS 2013.* 2013.

[53] P Sandborn. “Autonomous Maintenance for Through-Life Engineering”. In: *Through-life Engineering Services: Motivation, Theory, and Practice* (2015), pp. 341–357.

[54] Hagen Schempf et al. “GRISLEE: Gasmain repair and inspection system for live entry environments”. In: *International Journal of Robotics Research.* 2003.

[55] Joachim Hertzberg et al. “Landmark-based autonomous navigation in sewerage pipes”. In: *Proceedings of the 1st Euromicro Workshop on Advanced Mobile Robots, EUROBOT 1996.* 1996.

[56] Amir A.F. Nassiraei et al. “Concept and design of a fully autonomous sewer pipe inspection mobile robot ”KANTARO””. In: *Proceedings - IEEE International Conference on Robotics and Automation.* 2007.

[57] Bo Liu et al. “Eddy current array instrument and probe for crack detection of aircraft tubes”. In: *2010 International Conference on Intelligent Computation Technology and Automation, ICICTA 2010*. 2010.

[58] Markus Eich et al. “Design and control of a lightweight magnetic climbing robot for vessel inspection”. In: *2011 19th Mediterranean Conference on Control and Automation, MED 2011*. 2011.

[59] Daehie Hong et al. “Tethered Mobile Robot for Automating Highway Maintenance Operations”. In: *Robotics and Computer-Integrated Manufacturing* (1997).

[60] María Urdaneta et al. “Development of a novel autonomous robot for navigation and: Inspect in oil wells”. In: *Control Engineering and Applied Informatics* (2012).

[61] Daniel Schmidt et al. “Climbing robots for maintenance and inspections of vertical structures - A survey of design aspects and technologies”. In: *Robotics and Autonomous Systems* (2013).

[62] Akihiko Nagakubo et al. “Walking and running of the quadruped wall-climbing robot”. In: *Proceedings - IEEE International Conference on Robotics and Automation*. 1994.

[63] Bing L. Luk et al. “Intelligent legged climbing service robot for remote maintenance applications in hazardous environments”. In: *Robotics and Autonomous Systems* (2005).

[64] J. López et al. “WatchBot: A building maintenance and surveillance system based on autonomous robots”. In: *Robotics and Autonomous Systems* (2013).

[65] P W Prickett. “An integrated approach to autonomous maintenance management”. In: *Integrated Manufacturing Systems* 10.4 (1999), pp. 233–243.

[66] Giuseppe Curcurú et al. *Bayesian approach in the predictive maintenance policy*. Tech. rep.

[67] L. Schegner L. Urbas M. Krauss J. Birk. “Autonomie und Assistenz in der Prozessindustrie”. In: *Tagungsband Automation* (2018), pp. 519–528.

[68] Thomas Gamer et al. “The autonomous industrial plant – future of process engineering, operations and maintenance”. In: *Journal of Process Control* 88 (2020), pp. 101–110.

[69] Jim Davis et al. “Smart manufacturing, manufacturing intelligence and demand-dynamic performance”. In: *Computers and Chemical Engineering* (2012).

[70] Defang Li. “Perspective for smart factory in petrochemical industry”. In: *Computers and Chemical Engineering* (2016).

[71] Ian David Lockhart Bogle. “A Perspective on Smart Process Manufacturing Research Challenges for Process Systems Engineers”. In: *Engineering* (2017).

[72] Gregory H. Watson. “Digital hammers and electronic nails - Tools of the next generation”. In: *Quality Progress* (1998).

[73] K. Bowers et al. “Vox Populi 4.0: big data tools zoom in on the voice of the customer”. In: *Quality Progress* (), pp. 32–39.

[74] N. Radziwill. “Designing a Quality 4.0 Strategy and Selecting High-Impact Initiatives”. In: *ASQ Quality 4.0 Summit, Disruption, Innovation and Change.Dallas: ASQ*. 2018.

[75] Yann Lecun et al. *Deep learning*. 2015.

[76] Ross Girshick. “Fast R-CNN”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015.

[77] Jinjiang Wang et al. “Deep learning for smart manufacturing: Methods and applications”. en. In: *Journal of Manufacturing Systems* 48 (July 2018), pp. 144–156.

[78] Tao Zhang et al. “Current trends in the development of intelligent unmanned autonomous systems”. In: *Frontiers of Information Technology and Electronic Engineering* 18.1 (2017), pp. 68–85.

[79] Ross Girshick et al. “Rich feature hierarchies for accurate object detection and semantic segmentation”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2014.

[80] Alex Krizhevsky et al. “ImageNet classification with deep convolutional neural networks”. In: *Communications of the ACM* (2017).

[81] Po Sen Huang et al. “Learning deep structured semantic models for web search using clickthrough data”. In: *International Conference on Information and Knowledge Management, Proceedings*. 2013.

[82] Ivan Tritchkov et al. “Verification & validation of decentralized, self-organizing cyber-physical production systems: A blueprint process for testing cyber-physical production systems with self- properties”. In: *Proceedings - IEEE 1st International Workshops on Foundations and Applications of Self-Systems, FAS-W 2016*. Institute of Electrical and Electronics Engineers Inc., Dec. 2016, pp. 112–117.

[83] Paulo Leitão et al. “ADACOR: A Collaborative Production Automation and Control Architecture”. In: *Intelligent Systems, IEEE* 20 (Feb. 2005), pp. 58–66.

[84] Charlotta Johnsson. “ISA 95 - how and where can it be applied”. In: Oct. 2004.

[85] Yan Lu et al. *Current Standards Landscape for Smart Manufacturing Systems*. 2016.

[86] P. J. Antsaklis et al. “Towards intelligent autonomous control systems: Architecture and fundamental issues”. In: *Journal of Intelligent and Robotic Systems* (1989).

[87] Paulo Leitão et al. "Towards autonomy, self-organisation and learning in holonic manufacturing". In: *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*. Vol. 2691. Springer Verlag, 2003, pp. 544–553.

[88] Eddy Bajic et al. "Holonic Manufacturing with Intelligent Objects". In: *Knowledge and Technology Integration in Production and Services* (2002), pp. 297–304.

[89] Hong-Seok Park et al. "Autonomy for Smart Manufacturing". In: *Journal of the Korean Society for Precision Engineering* 31.4 (Apr. 2014), pp. 287–295.

[90] J Albus et al. "Autonomy in Engineering Systems: What is it and Why is it Important? Setting the Stage: Some Autonomous Thoughts on Autonomy". In: October 1998 (2005), pp. 520–521.

[91] H. S. Park. *From Automation to Autonomy - a New Trend for Smart Manufacturing*. 2013, pp. 075–110.

[92] Filippo Santoni de Sio et al. "Meaningful human control over autonomous systems: A philosophical account". In: *Frontiers Robotics AI* 5.FEB (2018), pp. 1–14.

[93] Saeid Nahavandi. "Trusted Autonomy Between Humans and Robots: Toward Human-on-the-Loop in Robotics and Autonomous Systems". In: *IEEE Systems, Man, and Cybernetics Magazine* 3.1 (2017), pp. 10–17.

[94] David Danks et al. "Algorithmic bias in autonomous systems". In: *IJCAI International Joint Conference on Artificial Intelligence* 0.August 2017 (2017), pp. 4691–4697.

[95] Robert Alexander et al. "Safety Lifecycle Activities for Autonomous Systems Development". In: *4th SEAS DTC Technical Conference - Edinburgh 2010* 106.11 (2010), pp. 1323–1330.

[96] R D Alexander et al. "Structuring safety cases for autonomous systems". In: *IET Conference Publications* 542 CP (2008).

[97] Miriam Gil et al. "Designing human-in-the-loop autonomous Cyber-Physical Systems". In: *International Journal of Human Computer Studies* 130.June 2018 (2019), pp. 21–39.

[98] Kush R. Varshney et al. "On the Safety of Machine Learning: Cyber-Physical Systems, Decision Sciences, and Data Products". In: *Big Data* 5.3 (Sept. 2017), pp. 246–255.

[99] Sebin Jung et al. "Exploring Challenges in Developing a Smart and Effective Assistive System for Improving the Experience of the Elderly Drivers". In: *Chinese Journal of Mechanical Engineering (English Edition)* 30.5 (Sept. 2017), pp. 1133–1149.

[100] Michael A. Goodrich et al. "Human-robot interaction: A survey". In: *Foundations and Trends in Human-Computer Interaction* 1.3 (2007), pp. 203–275.

[101] R. D. Alexander et al. “The role of the human in an autonomous system”. In: *IET Conference Publications* 2009.555 CP (2009).

[102] Marie Pierre Pacaux-Lemoine et al. “Towards human-based industrial cyber-physical systems”. In: *Proceedings - 2018 IEEE Industrial Cyber-Physical Systems, ICPS 2018* (2018), pp. 615–620.

[103] Yi Chen et al. “Intelligent autonomous pollination for future farming - A micro air vehicle conceptual framework with artificial intelligence and human-in-the-loop”. In: *IEEE Access* 7 (2019), pp. 119706–119717.

[104] Rino Falcone et al. “The human in the loop of a delegated agent: The theory of adjustable social autonomy”. In: *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.* 31.5 (2001), pp. 406–418.

[105] Gunar Schirner et al. “The Future of Human-in-the-Loop Cyber-Physical Systems”. In: (2013).

[106] Lu Feng et al. “Synthesis of Human-in-the-Loop Control Protocols for Autonomous Systems”. In: *IEEE Transactions on Automation Science and Engineering* 13.2 (2016), pp. 450–462.

[107] Kerstin Dautenhahn. “The art of designing socially intelligent agents: science, fiction, and the human in the loop”. In: *Applied Artificial Intelligence* 12.7-8 (1998), pp. 573–617.

[108] Luka Peternel et al. “Robotic assembly solution by human-in-the-loop teaching method based on real-time stiffness modulation”. In: *Autonomous Robots* 42.1 (2018), pp. 1–17.

[109] Andreas Birk et al. “Rescue robotics - A crucial milestone on the road to autonomous systems”. In: *Advanced Robotics* 20.5 (2006), pp. 595–605.

[110] Paulo Leitão. “Agent-based distributed manufacturing control: A state-of-the-art survey”. In: *Engineering Applications of Artificial Intelligence* 22.7 (Oct. 2009), pp. 979–991.

[111] Michael Hülsmann et al. *Understanding Autonomous Cooperation and Control in Logistics.* Springer, 2007.

[112] Emiliano Sisinni et al. “Industrial internet of things: Challenges, opportunities, and directions”. In: *IEEE Transactions on Industrial Informatics* 14.11 (Nov. 2018), pp. 4724–4734.

[113] Jiangfeng Cheng et al. “Industrial IoT in 5G environment towards smart manufacturing”. In: *Journal of Industrial Information Integration* 10 (June 2018), pp. 10–19.

[114] International Electrotechnical Commision (IEC). *IoT 2020: Smart and secure IoT platform.* Tech. rep. Switzerland, 2016.

[115] Guodong Shao et al. “Digital Twin for Smart Manufacturing: The Simulation Aspect”. In: *Proceedings - Winter Simulation Conference* 2019-Decem.Bolton 2016 (2019), pp. 2085–2098.

[116] James Moyne et al. “A Requirements Driven Digital Twin Framework: Specification and Opportunities”. In: *IEEE Access* 8 (2020), pp. 107781–107801.

[117] Luis Martins et al. “Literature review on autonomous production control methods”. In: *Enterprise Information Systems* 00.00 (2020), pp. 1–13.

[118] Sebastian Grundstein et al. “A new method for autonomous control of complex job shops – Integrating order release, sequencing and capacity control to meet due dates”. In: *Journal of Manufacturing Systems* 42 (Jan. 2017), pp. 11–28.

[119] Kosmas Alexopoulos et al. “Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing”. In: *International Journal of Computer Integrated Manufacturing* 33.5 (2020), pp. 429–439.

[120] Carlos Toro et al. “A perspective on knowledge based and intelligent systems implementation in industrie 4.0”. In: *Procedia Computer Science* 60.1 (2015), pp. 362–370.

[121] Xifan Yao et al. “From Intelligent Manufacturing to Smart Manufacturing for Industry 4.0 Driven by Next Generation Artificial Intelligence and Further on”. In: *Proceedings - 2017 5th International Conference on Enterprise Systems: Industrial Digitalization by Enterprise Systems, ES 2017*. Institute of Electrical and Electronics Engineers Inc., Nov. 2017, pp. 311–318.

[122] László Monostori. “AI and machine learning techniques for managing complexity, changes and uncertainties in manufacturing”. In: *Engineering Applications of Artificial Intelligence* 16.4 (2003), pp. 277–291.

[123] L. Monostori et al. “Agent-based systems for manufacturing”. In: *CIRP Annals - Manufacturing Technology* 55.2 (2006), pp. 697–720.

[124] Alexander Bannat et al. “Artificial cognition in production systems”. In: *IEEE Transactions on Automation Science and Engineering* 8.1 (2011), pp. 148–174.

[125] Louise A. Dennis et al. “Verifiable Self-Aware Agent-Based Autonomous Systems”. In: *Proceedings of the IEEE* 108.7 (May 2020), pp. 1011–1026.

[126] Shiyong Wang et al. “Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination”. In: *Computer Networks* 101 (June 2016), pp. 158–168.

[127] Hong Seok Park et al. “An autonomous manufacturing system for adapting to disturbances”. In: *International Journal of Advanced Manufacturing Technology* 56.9-12 (Oct. 2011), pp. 1159–1165.

[128] Andre Lebioda et al. *Control of Cyber-Physical Production Systems: A Concept to Increase the Trustworthiness within Multi-Agent Systems with Distributed Ledger Technology Completed Research Paper*. Tech. rep. 2019.

[129] Stefan Bussmann. "Agent-Based Control Systems". In: *IEEE Control Systems* 23.3 (2003), pp. 61–73.

[130] Harley Oliff et al. "Reinforcement learning for facilitating human-robot-interaction in manufacturing". In: *Journal of Manufacturing Systems* 56.June (2020), pp. 326–340.

[131] Arndt Lüder et al. "Design pattern for agent based production system control - A survey". In: *IEEE International Conference on Automation Science and Engineering* 2017-Augus (2017), pp. 717–722.

[132] WARRENDALE. "https://www.sae.org/news/press-room/2018/12". In: *SAE International* (2018).

[133] Bernhard Langefeld. "Rise of the machines – How robots and artificial intelligence are shaping the future of autonomous production". In: (), pp. 3–10.

[134] Dennis Bauer et al. "Characterization of Autonomous Production by a Stage Model". In: *Procedia CIRP* 81 (2019), pp. 192–197.

[135] Fritz Zwicky et al. *New methods of thought and procedure: Contributions to the symposium on methodologies*. Springer Science & Business Media, 2012.

[136] Fritz Zwicky. "Discovery, invention, research through the morphological approach". In: (1969).

[137] T Ritchey. "General Morphological Analysis: A General Method for Non-Quantifiable Modeling". In: *Conference for Operational Analysis, Brussels*. 2002.

[138] Philipp Gölzer et al. "Data Processing Requirements of Industry 4 . 0 –". In: *European Conference on Information Systems ECIS* 0.March 2019 (2015), pp. 1–13.

[139] Laura Kassner et al. "M2DDM – A Maturity Model for Data-Driven Manufacturing". In: 63 (2017), pp. 173–178.

[140] Petter Gottschalk. "Maturity levels for interoperability in digital government". In: *Government Information Quarterly* 26.1 (Jan. 2009), pp. 75–81.

[141] David Chen et al. "Architectures for enterprise integration and interoperability: Past, present and future". In: *Computers in Industry* 59.7 (Sept. 2008), pp. 647–659.

[142] Stanislav Chankov et al. "Synchronization in manufacturing systems: quantification and relation to logistics performance". In: *International Journal of Production Research* (2016).

[143] Mohamed Anis Dhuieb et al. "Context-awareness: A Key Enabler for Ubiquitous Access to Manufacturing Knowledge". In: *Procedia CIRP* 41 (2016), pp. 484–489.

[144] Patrick Rosenberger et al. “Context-awareness in industrial applications: Definition, classification and use case”. In: *Procedia CIRP* 72 (2018), pp. 1172–1177.

[145] Sebastian Scholze et al. “Context Awareness for Flexible Manufacturing Systems Using Cyber Physical Approaches”. In: 2016, pp. 107–115.

[146] Dražen Nadoveza et al. “Concept for Context-Aware Manufacturing Dashboard Applications”. In: *IFAC Proceedings Volumes* 46.9 (2013), pp. 204–209.

[147] Lamia Zouhaier et al. “Context Awareness Systems: Architecture and Context Modeling”. In: *Internatinal conference on Control, Engineering & Information Technology* 2.December (2013), pp. 130–135.

[148] L. Wang et al. “Symbiotic human-robot collaborative assembly”. In: *CIRP Annals* 68.2 (Jan. 2019), pp. 701–726.

[149] David Romero et al. *The Operator 4.0: Towards socially sustainable factories of the future*. Jan. 2020.

[150] Shirine El Zaatari et al. *Cobot programming for collaborative industrial tasks: An overview*. June 2019.

[151] Gudela Grote. “Safety and autonomy: A contradiction forever?” In: *Safety Science* 127 (July 2020).

[152] Christian Schröder. “The challenges of industry 4.0 for small and medium-sized enterprises”. In: *Friedrich-Ebert-Stiftung: Bonn, Germany* (2016).

[153] Brian Buntz. “https://www.iotworldtoday.com/2016/04/20/top-10-reasons-people-aren-t-embracing-iot/”. In: *IOT world today* (2016).

[154] Engelbert Westkämper et al. *Strategien der Produktion*. Springer, 2016.

[155] Chun-Fai Chan et al. “Detecting Anomalies in Programmable Logic Controllers Using Unsupervised Machine Learning”. In: *IFIP International Conference on Digital Forensics*. Springer. 2019, pp. 119–130.

[156] Paul Horn. “Autonomic computing: IBM’s perspective on the state of information technology”. In: (2001).

[157] Shenin Hassan et al. “Autonomic Computing Paradigm to Support System’s Development”. en. In: *2009 Second International Conference on Developments in eSystems Engineering*. Abu Dhabi, UAE: IEEE, Dec. 2009, pp. 273–278.

[158] J.O. Kephart et al. “The vision of autonomic computing”. en. In: *Computer* 36.1 (Jan. 2003), pp. 41–50.

[159] Mazeiar Salehie et al. “Autonomic Computing: Emerging Trends and Open Problems”. en. In: (2005), p. 7.

[160] M.G. Hinchey et al. “Self-Managing Software”. en. In: *Computer* 39.2 (Feb. 2006), pp. 107–109.

[161] R. Sterritt et al. “Autonomic Computing - a means of achieving dependability?” en. In: *10th IEEE International Conference and Workshop on the Engineering of Computer-Based Systems, 2003. Proceedings.* Huntsville, AL, USA: IEEE Comput. Soc, 2003, pp. 247–251.

[162] Daniel Stock et al. “System Architectures for Cyber-Physical Production Systems enabling Self-X and Autonomy”. en. In: *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA).* Vienna, Austria: IEEE, Sept. 2020, pp. 148–155.

[163] Jarkko Ruonala. *Understanding Measurement Performance and Specifications.* `https://www.vaisala.com/en/blog/2020-07/understanding-measurement-performance-and-specifications`. 2016.

[164] Jacob Yerushalmy. “Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques”. In: *Public Health Reports (1896-1970)* (1947), pp. 1432–1449.

[165] Shen Yin et al. “Big data for modern industry: challenges and trends [point of view]”. In: *Proceedings of the IEEE* 103.2 (2015), pp. 143–146.

[166] Chongzhen Zhang et al. “When Autonomous Systems Meet Accuracy and Transferability through AI: A Survey”. In: *Patterns* 1.4 (July 2020), p. 100050.

[167] Monika Hengstler et al. “Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices”. In: *Technological Forecasting and Social Change* 105 (2016), pp. 105–120.

[168] Alan FT Winfield et al. “Ethical governance is essential to building trust in robotics and artificial intelligence systems”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376.2133 (2018), p. 20180085.

[169] Ifeyinwa Angela Ajah et al. “Exploring the benefits of the 4th industrial revolution: The Nigerian experience”. In: *AFRREV STECH: An International Journal of Science and Technology* 8.1 (Aug. 2019), p. 22.

[170] Resul Kurt. “Industry 4.0 in Terms of Industrial Relations and Its Impacts on Labour Life”. In: *Procedia Computer Science* 158 (2019), pp. 590–601.

[171] Elena G. Popkova et al., eds. *Industry 4.0: Industrial Revolution of the 21st Century.* Vol. 169. Studies in Systems, Decision and Control. Cham: Springer International Publishing, 2019.

[172] Jeremy Greenwood. *The third industrial revolution.* Washington, D.C., DC: AEI Press, 1997, pp. 1–6.

[173] Alla Kravchenko et al. “The Forth Industrial Revolution: New Paradigm of Society Development or Posthumanist Manifesto”. In: *Philosophy and Cosmology* 22 (Jan. 2019), pp. 120–128.

[174] Thorsten Wuest et al. “Machine learning in manufacturing: advantages, challenges, and applications”. In: *Production & Manufacturing Research* 4.1 (Jan. 2016), pp. 23–45.

[175] Maria Flavia Mogos et al. “Enablers and inhibitors of Industry 4.0: results from a survey of industrial companies in Norway”. In: *Procedia CIRP* 81 (2019), pp. 624–629.

[176] J Chaplin et al. “A five-stage model of autonomy and its features: literature review in manufacturing context”. DiManD Deliverable 4.1a. 2021.

[177] Bernhard Langefeld et al. *Rise of the machines - How robots and arti cial intelligence are shaping the future of autonomous production.* Tech. rep. Munich, Germany, 2019.

[178] Chetna Chauhan et al. “Barriers to industry 4.0 adoption and its performance implications: An empirical investigation of emerging economy”. In: *Journal of Cleaner Production* (Oct. 2020), p. 124809.

[179] Rupert Glass et al. “Identifying the barriers to Industrie 4.0”. In: *Procedia CIRP* 72 (2018), pp. 985–988.

[180] Dóra Horváth et al. “Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities?” In: *Technological Forecasting and Social Change* 146.May (2019), pp. 119–132.

[181] Ron Davies. *Industry 4.0 Digitalisation for productivity and growth.* Tech. rep. European Parlamentary Research Service, 2015.

[182] Automation Alley. *Industry 4.0 is Here. Are We Ready?* Tech. rep. Troy, MI (USA): Automation Alley, 2017, pp. 1–15.

[183] Siri Adolph et al. “Challenges and approaches to competency development for future production”. In: *Journal of International Scientific Publications–Educational Alternatives* 12.1 (2014), pp. 1001–1010.

[184] Wilhelm Bauer et al. “Transforming to a hyper-connected society and economy–towards an “Industry 4.0””. In: *Procedia Manufacturing* 3 (2015), pp. 417–424.

[185] Daniel Kiel et al. “The influence of the Industrial Internet of Things on business models of established manufacturing companies–A business level perspective”. In: *Technovation* 68 (2017), pp. 4–19.

[186] Judit Nagy. “Az Ipar 4.0 fogalma és kritikus kérdései–vállalati interjúk alapján”. In: *Vezetéstudomány-Budapest Management Review* 50.1 (2019), pp. 14–26.

[187] Industry McKinsey. “4.0 after the initial hype”. In: *Where manufacturers are finding value and how they can best capture it. McKinsey Digital* (2016).

[188] David Danks et al. *Algorithmic Bias in Autonomous Systems.* Tech. rep. 2017.

[189] Rui Ribeiro. "Digital Transformation: The Evolution of the Enterprise Value Chains". In: Springer, Singapore, Feb. 2021, pp. 290–302.

[190] Stephen Jia Wang et al. "Big data for urban sustainability: A human-centered perspective". In: *Big Data for Urban Sustainability: A Human-Centered Perspective* (2018), pp. 1–160.

[191] Daniel Stock et al. "System Architectures for Cyber-Physical Production Systems enabling Self-X and Autonomy". In: *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA* 2020-Septe (2020), pp. 148–155.

[192] Gerlind Wisskirchen et al. "Artificial Intelligence and Robotics and Their Impact on the Workplace". In: *IBA Global Employment Institute* April (2017), p. 120.

[193] Roberto V Zicari. "Big data: Challenges and opportunities". In: *Big data computing* 564 (2014), p. 103.

[194] ISO/IEC. *ISO 16100-6:2011(E): Industrial Automation Systems and Integration—Manufacturing Software Capability Profiling for Interoperability*. 2015.

[195] Sharon J Kemmerer. *Manufacturing Interoperability Program, a Synopsis NISTIR 7533*. Tech. rep.

[196] Ron Davies. *Industry 4.0 Digitalisation for productivity and growth*. Tech. rep. European Parliament, 2015.

[197] C. Y. Park et al. "Predictive situation awareness model for smart manufacturing". In: *2017 20th International Conference on Information Fusion (Fusion)*. 2017, pp. 1–8.

[198] Sunpreet Singh et al. "Towards zero waste manufacturing: A multidisciplinary review". In: *Journal of cleaner production* 168 (2017), pp. 1230–1243.

[199] Xuan Luu Hoang et al. "Systematization approach for the adaptation of manufacturing machines". In: *2016 IEEE 21st international conference on emerging technologies and factory automation (ETFA)*. IEEE. 2016, pp. 1–4.

[200] Joseph Berk. *Cost reduction and optimization for manufacturing and industrial companies*. Vol. 2. John Wiley & Sons, 2010.

[201] James C Wetherbe et al. "Zero based budgeting in the planning process". In: *Strategic Management Journal* 2.1 (1981), pp. 1–14.

[202] John Raju. "A conceptual design and cost optimisation methodology". In: *44th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. 2003, p. 1505.

[203] Jeffrey P Kharoufeh et al. "Reliability of manufacturing equipment in complex environments". In: *Annals of Operations Research* 209.1 (2013), pp. 231–254.

[204] Guanling Chen et al. *A Survey of Context-Aware Mobile Computing Research*. Tech. rep. USA, 2000.

[205] Juergen Lenz et al. "Data-driven Context Awareness of Smart Products in Discrete Smart Manufacturing Systems". In: *Procedia Manufacturing* 52 (2020), pp. 38–43.

[206] Loan Thi Ngoc Dinh et al. "A survey on context awareness in big data analytics for business applications". In: *Knowledge and Information Systems* 62.9 (Sept. 2020), pp. 3387–3415.

[207] Kalle J. Lyytinen et al. "Surfing the Next Wave: Design and Implementation Challenges of Ubiquitous Computing". In: *Communications of the Association for Information Systems* 13 (2004).

[208] Marie Pierre Pacaux-Lemoine et al. "Designing intelligent manufacturing systems through Human-Machine Cooperation principles: A human-centered approach". In: *Computers and Industrial Engineering* 111 (Sept. 2017), pp. 581–595.

[209] Damien Trentesaux et al. "A human-centred design to break the myth of the "magic human" in intelligent manufacturing systems". In: *Studies in Computational Intelligence*. Vol. 640. Springer Verlag, 2016, pp. 103–113.

[210] Paola Fantini et al. "Placing the operator at the centre of Industry 4.0 design: Modelling and assessing human activities within cyber-physical systems". In: *Computers and Industrial Engineering* 139 (Jan. 2020).

[211] John O. Oyekan et al. "The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans". In: *Robotics and Computer-Integrated Manufacturing* 55 (Feb. 2019), pp. 41–54.

[212] Jeremy A. Marvel et al. "Towards effective interface designs for collaborative HRI in manufacturing". In: *ACM Transactions on Human-Robot Interaction* 9.4 (Oct. 2020).

[213] Costanza Messeri et al. "On the effects of leader-follower roles in dyadic human-robot synchronisation". In: *IEEE Transactions on Cognitive and Developmental Systems* (2020).

[214] J. Douglas Orton et al. "Loosely Coupled Systems: A Reconceptualization". In: *The Academy of Management Review* 15.2 (Apr. 1990), p. 203.

[215] Maksim Jenihhin et al. "Challenges of reliability assessment and enhancement in autonomous systems". In: *2019 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)*. IEEE. 2019, pp. 1–6.

[216] John McDermid et al. "Towards a Framework for Safety Assurance of Autonomous Systems". In: *Artificial Intelligence Safety* (2019).

[217] Silviya Serafimova. "Whose morality? Which rationality? Challenging artificial intelligence as a remedy for the lack of moral enhancement". In: *Humanities and Social Sciences Communications* 7.1 (Dec. 2020).

[218] Thomas M Powers. “Prospects for a Kantian Machine”. In: *IEEE Intelligent Systems* (2006), pp. 46–51.

[219] Michael Anderson et al. “Machine Ethics: Creating an Ethical Intelligent Agent”. In: *AI Magazine* (2007).

[220] Thomas M. Powers. *Philosophy and computing: essays in epistemology, philosophy of mind, logic, and ethics*. Ed. by Thomas M. Powers. Vol. 128. Cham: Springer International Publishing, 2017.

[221] “Top 3 Barriers to Industrial IoT Success and How to Overcome Them”. In: ().

[222] Dan Craigen et al. “Defining cybersecurity”. In: *Technology Innovation Management Review* 4.10 (2014).

[223] Julian Jang-Jaccard et al. “A survey of emerging threats in cybersecurity”. In: *Journal of Computer and System Sciences* 80.5 (2014), pp. 973–993.

[224] Priscilla Koepke. “Cybersecurity information sharing incentives and barriers”. In: *Sloan School of Management at MIT University* (2017).

[225] Sundeep Bhandari et al. *Reproducibility in research: the role of measurement science workshop Recommendations for the infrastructure underpinning digital transformation.* Teddington, Middlesex: National Physical Laboratory, 2020, p. 14.

[226] Tobe Toben. “A Formal Model of Reliable Sensor Perception”. In: *Smart Sensing and Context*. 2010, pp. 94–107.

[227] Goran Vasiljević et al. “High-accuracy vehicle localization for autonomous warehousing”. In: *Robotics and Computer-Integrated Manufacturing* 42 (Dec. 2016), pp. 1–16.

[228] Claudio Bettini et al. “A survey of context modelling and reasoning techniques”. In: *Pervasive and Mobile Computing* 6.2 (Apr. 2010), pp. 161–180.

[229] Tobe Toben et al. “Safety and Precision of Spatial Context Models for Autonomous Systems”. In: *Electronic Notes in Theoretical Computer Science* 297 (Dec. 2013), pp. 75–88.

[230] Robert J. Baron. “Mechanisms of human facial recognition”. In: *International Journal of Man-Machine Studies* 15.2 (Aug. 1981), pp. 137–178.

[231] Benjamin Balaguer et al. “Evaluating maps produced by urban search and rescue robots: lessons learned from RoboCup”. In: *Autonomous Robots* 27.4 (Nov. 2009), pp. 449–464.

[232] Davide Ronzoni et al. “AGV global localization using indistinguishable artificial landmarks”. In: *2011 IEEE International Conference on Robotics and Automation*. IEEE, May 2011, pp. 287–292.

[233] Raffaello D'Andrea. "Guest Editorial: A Revolution in the Warehouse: A Retrospective on Kiva Systems and the Grand Challenges Ahead". In: *IEEE Transactions on Automation Science and Engineering* 9.4 (Oct. 2012), pp. 638–639.

[234] Jürgen Sturm et al. "A benchmark for the evaluation of RGB-D SLAM systems". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2012, pp. 573–580.

[235] Christoph Sprunk et al. "An experimental protocol for benchmarking robotic indoor navigation". In: *Experimental Robotics*. Springer. 2016, pp. 487–504.

[236] Sebastian Thrun et al. *Probabilistic robotics, vol. 1*. 2005.

[237] T. von Leipzig et al. "Initialising Customer-orientated Digital Transformation in Enterprises". In: *Procedia Manufacturing* 8 (2017).

[238] Royal Academy of Engineering (Great Britain). *Autonomous systems : social, legal and ethical issues*. Tech. rep. 2009, p. 15.

[239] Shervin Shahrdar et al. "A survey on trust in autonomous systems". In: *Advances in Intelligent Systems and Computing*. Vol. 857. Springer Verlag, 2019, pp. 368–386.

[240] Andreas Kaplan et al. "Rulers of the world, unite! The challenges and opportunities of artificial intelligence". In: *Business Horizons* (2019).

[241] Eva Fraedrich et al. "Societal and individual acceptance of autonomous driving". In: *Autonomous Driving: Technical, Legal and Social Aspects*. Springer Berlin Heidelberg, Jan. 2016, pp. 621–640.

[242] Jihye Lee et al. "Influencing Factors on Social Acceptance of Autonomous Vehicles and Policy Implications". In: *Proceedings of PICMET '18: Technology Management for Interconnected World* (2018).

[243] A. M. TURING. "I.—COMPUTING MACHINERY AND INTELLIGENCE". In: *Mind* LIX.236 (Oct. 1950).

[244] Philippa Foot. *The Problem of Abortion and the Doctrine of the Double Effect*. Tech. rep. 1967.

[245] Lucia Vesnic-Alujevic et al. "Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks". In: *Telecommunications Policy* 44.6 (July 2020).

[246] Gianmarco Veruggio. "Roboethics". In: *IEEE Robotics and Automation Magazine* 17.2 (June 2010).

[247] Gianmarco Veruggio et al. *Roboethics: Ethics applied to robotics*. Mar. 2011.

[248] Gianmarco Veruggio et al. "Roboethics: Social and Ethical Implications". In: *Springer Handbook of Robotics*. Cham: Springer International Publishing, 2016.

[249] Luciano Floridi et al. “A Unified Framework of Five Principles for AI in Society”. In: *Harvard Data Science Review* (June 2019).

[250] Smriti Parsheera. *A gendered perspective on artificial intelligence.* Tech. rep. 2018.

[251] Sarah Bird et al. “Fairness-aware machine learning: Practical challenges and lessons learned”. In: *The Web Conference 2019 - Companion of the World Wide Web Conference, WWW 2019*. Association for Computing Machinery, Inc, May 2019, pp. 1297–1298.

[252] European Group on Ethics in Science and New Technologies. *Statement on Artificial Intelligence, Robotics and ’Autonomous’ Systems*. Tech. rep. Brussels: European Commission, 2018.

[253] Susanne Schmetkamp. “Understanding A.I. — Can and Should we Empathize with Robots?” In: *Review of Philosophy and Psychology* (Dec. 2020).

[254] Robert M. Veatch. *Is there a common morality?* 2003.

[255] David Steinberg. *The Multidisciplinary Nature of Morality and Applied Ethics*. Cham: Springer International Publishing, 2020.

[256] Aaron M. Johnson et al. “THE MORALITY OF AUTONOMOUS ROBOTS”. In: *Journal of Military Ethics* 12.2 (July 2013).

[257] Salvador Cervantes et al. “Toward ethical cognitive architectures for the development of artificial moral agents”. In: *Cognitive Systems Research* 64 (Dec. 2020), pp. 117–125.

[258] John Danaher. “The rise of the robots and the crisis of moral patiency”. In: *AI and Society* 34.1 (Mar. 2019), pp. 129–136.

[259] Miranda Bogen et al. “Awareness in practice: Tensions in access to sensitive attribute data for antidiscrimination”. In: *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, Inc, Jan. 2020, pp. 492–500.

[260] Rediet Abebe et al. “Roles for computing in social change”. In: *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, Inc, Jan. 2020, pp. 252–260.

[261] Sara Gerke et al. “Ethical and legal challenges of artificial intelligence-driven healthcare”. In: *Artificial Intelligence in Healthcare*. Elsevier, 2020, pp. 295–336.

[262] Alex Hanna et al. “Towards a critical race methodology in algorithmic fairness”. In: *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, Inc, Jan. 2020, pp. 501–512.

[263] *AI software defines people as male or female. That’s a problem - CNN.*

[264] Eun Seo Jo et al. “Lessons from archives: Strategies for collecting sociocultural data in machine learning”. In: *FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.* Association for Computing Machinery, Inc, Jan. 2020, pp. 306–316.

[265] Kenneth Holstein et al. “Improving fairness in machine learning systems: What do industry practitioners need?” In: *Conference on Human Factors in Computing Systems - Proceedings.* Association for Computing Machinery, May 2019.

[266] T Velnampy et al. “An assocaition between organisational growth and profitability: A study of commercial bank of Ceylon Ltd Srilanka”. In: *Universitatii Bucuresti. Analele. Seria Stiinte Economice si Administrative* 2 (2008), p. 43.

[267] Gustav Ranis et al. “A theory of economic development”. In: *The American economic review* (1961), pp. 533–565.

[268] Alok Raj et al. “Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective”. In: *International Journal of Production Economics* 224 (June 2020).

[269] Antonella Petrillo et al. “Fourth Industrial Revolution: Current Practices, Challenges, and Opportunities”. In: Feb. 2018.

[270] P. Macurová et al. “The driving factors, risks and barriers of the industry 4.0 concept”. In: *Journal of Applied Economic Sciences* 12 (Jan. 2017), pp. 2003–2011.

[271] Ayhan Görmüş. “Future of Work with the Industry 4.0”. In: Oct. 2019, pp. 317–323.

[272] Aidan Fuller et al. “Digital Twin: Enabling Technologies, Challenges and Open Research”. In: (May 2020).

[273] Michael E Porter et al. “Toward a New Conception of the Environment-Competitiveness Relationship”. In: *Journal of Economic Perspectives* 9.4 (Nov. 1995), pp. 97–118.

[274] Katherine L. Christ et al. “Material flow cost accounting: a review and agenda for future research”. In: *Journal of Cleaner Production* 108 (Dec. 2015), pp. 1378–1389.

[275] Dieter Spath et al. *Produktionsarbeit der Zukunft-Industrie 4.0.* Vol. 150. Fraunhofer Verlag Stuttgart, 2013.

[276] Ruth Stock-Homburg. “Zukunft der Arbeitswelt 2030 als Herausforderung des Personalmanagements”. In: *Handbuch Strategisches Personalmanagement.* Springer, 2013, pp. 603–629.

[277] Roger Burritt et al. “Industry 4.0 and environmental accounting: a new revolution?” In: *Asian Journal of Sustainability and Social Responsibility* 1.1 (Dec. 2016), pp. 23–38.

[278] “What are the main barriers to innovation manufacturers face?” In: ().

Chapter 4

Versions

D4.1 Requirements specification for Distributed Manufacturing Systems	
Version - Date	**Comments and Recommendations**
V0-2021-02-21	Structured First Version
V1-2021-08-05	Revised Second Version