WORKING PAPER

Technology Scenario
'Artificial Intelligence in Industrie 4.0'

Table of Contents

Preface

The use of artificial intelligence (AI) is regarded as the major force behind innovation today. Whether in the private, public, or economic sphere, the use of artificial intelligence has an impact on society as a whole. It is not a new topic, though. The development of the first ideas, principles, and technologies already began in the middle of the last century. At that time, however, technologies were limited by limited digital memory capacities, computer processing power, and network bandwidths. Today, artificial intelligence is already deployable across the board, bearing in mind the steady advancements in this field and the resource needs which it creates.

Artificial intelligence is understood in this context as a collection of methods that enable a computer to solve tasks that would require intelligence if they were solved by human beings. Although intelligence per se is difficult to define, it is generally agreed that something cannot be considered intelligent if it does not have the capacity to learn and the capacity to solve problems independently. Thus machine learning (ML) is one of the central sub-areas of artificial intelligence, but it is not the only one. Most success stories today in the area of artificial intelligence involve a combination of learning and problem-solving. But for machines to learn, they need data: either large quantities of data for one-time analytical purposes, or quantities of data for repetitive re-learning purposes, or streams of data from which learning is continuously taking place. On the basis of data and through skilful connections, machine learning is able to reduce complexities and detect events or patterns, which it then uses to explain events, make predictions, or enable actions to be taken – and it is able to do this without explicit programming in the form of the usual 'if-then' routines or without classic automation and control engineering. Digitalisation is steadily permeating all areas and is making it possible to provide large quantities of data in an automated fashion. This is the exact data needed by artificial intelligence to generate benefit.

There are high expectations for the use of artificial intelligence: the highly diverse areas where it is already being applied or where it could be applied, and the current focus on it in the political sphere, the scientific community, and in consumer circles. In addition to the technical issues involved with AI, it also raises issues of data sovereignty, the future of work, and ethical usage, to name a few.

The use of new technologies and AI facilitates innovation in industrial production. Two factors are essential for success in this context. The first is the need for motivated and qualified employees who have learned to deal competently with AI. The second is the need for a supportive environment which, in addition to fostering a secure and structured use of AI, also furthers its acceptance at both the national and international levels. The following paper provides a differentiated view of the impact of AI in industrial manufacturing – Industrie 4.0 – from the variety of perspectives within Plattform Industrie 4.0 and its working groups. Contributions to it were also made by the so-called Plattform Lernende Systeme (Germany's platform for artificial intelligence) and by the supporting research on the technology programme Smart Services Worlds of the Federal Ministry for Economic Affairs and Energy. The paper is meant to encourage readers to inform themselves about AI, to recognise chains of effects, and to reach their own conclusions.

The purpose of this paper is to describe the new possibilities opened up by AI and to categorise the various levels of autonomous automation processes. This categorisation describes the possible ways of designing industrial processes using AI – from manual to autonomous – and should not be understood as an assessment. It provides a basis for discussing the uses of AI and for making recommendations for future actions.

The working groups of Plattform Industrie 4.0 have produced other papers supplementary to this one. These address such things as the special requirements and calls for action in the legal area ('AI and Law in the Context of Industrie 4.0') and in relation to security issues ('Artificial Intelligence in the Security Aspects of Industrie 4.0').

Introduction

Artificial intelligence (AI) is considered one of the key technologies indispensable to the sustainment of Germany's economic strength.[1] In addition to the sphere of Industrie 4.0 processes, AI shows a great deal of potential for added value in the manufacturing and service industries. The current focus on product-oriented forms of production and manufacturing in Industrie 4.0 is being supplanted by concepts of a more solution- and customer-oriented nature. In the future, rigidly defined chains of production and value will be transformed into flexible and highly dynamic ecological manufacturing and service systems. They will facilitate fully individualised production on an order-by-order basis. With customised customer requirements as the starting point, production systems are to become autonomously organised and production and logistic strategies are to optimise themselves with the help of AI. Essential to the implementation of AI in industrial processes is an administration shell that structures and connects data sources, learning mechanisms, system boundaries, active actions, and human interventions. The structured and standardised features and interfaces of the network components of the model reference architecture Industrie 4.0 (RAMI 4.0) also provide a useful basis for integrating AI technologies into industrial production or engineering.

To better understand these processes, the opportunities and challenges surrounding the use of AI technologies in industrial work processes are discussed in conjunction with the activities of Working Group 2 – Technological and Application Scenarios – of the Plattform Industrie 4.0. Also contributing to such discussions is the project group 'AI for Industrie 4.0', a project that intersects all of the other working groups. An intense exchange of information takes place here with the Plattform Lernende Systeme, the KI Bundesverband e.V. (Federal Association for Artificial Intelligence), and with the supporting research projects of the technology programme Smart Services Worlds of the Federal Ministry for Economic Affairs and Energy.

From the point of view of industry, artificial intelligence technologies are to be understood as methods and procedures that enable technical systems to perceive their environments, process what they have perceived, solve problems independently, find new ways of solving problems, make decisions, and especially learn from experience in order to become more proficient at completing tasks and actions. The use of AI technologies is meant to increase the efficiency and effectiveness of industrial processes. Relevant to this are factors such as costs, speed, precision, or problem-solving beyond what human beings are capable of. The general idea is that higher levels of autonomy in industrial processes can only be attained through cognitive capacities and these can be supplied by AI technologies. Relative to the capacities and intensity of use of AI, the need for human intervention in the processes decreases. The following discussion is only concerned, however, with AI-based decision-making processes that take place within system boundaries defined by human beings and that operate with AI technologies monitored by human beings. This covers nearly all AI systems currently in use or to be used in the near future in industrial applications.

This paper will discuss the intensity of the use of AI technologies based on the level of autonomy of industrial processes, which depends on how independently the system can master complex situations in an automated and self-learning manner. References to this discussion are found in the technical literature, for example in articles on modern simulation experiments. These investigate whether the optimal level of autonomy of a production system can be ascertained through the use of an autonomy index that defines the relationship between autonomous process steps and the total number of process steps of an industrial process.[2] The work of project group 'AI for Industrie 4.0' is also aligned to the discussion on (autonomous) levels of self-driving cars.[3] The following presentation represents the current status of the discussion in the project group, the development of which will continue during the course of the ongoing project.

---

1   https://www.bmbf.de/files/180718%20Eckpunkte_KI-Strategie%20final%20Layout.pdf.

2   Gronau: Der angemessene Grad von Autonomie in Cyber-Physischen Produktionssystemen; Industrie 4.0 Management 34 (2018) 6.

3   SAE International: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, J3016_201806, Rev 2018-06-15.

Defining artificial intelligence

From a general point of view, artificial intelligence is a sub-discipline of computer science. It deals with methods and technologies that enable a computer to carry out those tasks that would require intelligence if they were carried out by human beings.[4] In the report on artificial intelligence prepared by the VDI (Association of German Engineers), a technical approach was taken to defining the term: 'Artificial intelligence supplements technical systems with the ability to process tasks independently and efficiently'.[5] But there is not yet any generally accepted, unambiguous, and exact definition of the term.[6] AI is often used to describe computer systems that complete tasks after they have been trained with large amounts of data and that afterwards, possibly together with other methods, make decisions derived from the data already known to them. Depending on the quality and the quantity of the training data, the AI system can execute what it considers the 'right' action. With the help of learning algorithms, AI systems can continue learning during ongoing operations, through which the trained models are optimised and the data- and knowledge-bases extended.[7]

The reason for the current hype around artificial intelligence lies in its core technological function, i.e. that it can, in principle, be used to implement every formalisation of human and rational thinking and acting, which means learning, planning, and problem-solving. Included here are image and voice recognition, knowledge acquisition, machine learning, cognitive grasping of and automation of logical inferences, and the planning and implementing of industrial automation processes. The artificial intelligence currently in use is human-oriented and focuses on using computers to support the activities of human beings through AI systems.

For the sake of simplifying the description of the possibilities of artificial intelligence, a difference is often drawn between 'weak' and 'strong' AI in the application of AI technologies.[8] Weak AI is understood as the mechanical ability to act in a way as if the machine were intelligent. Strong AI would be if the machine does in fact 'think', i.e. when thinking is not merely simulated. The philosophical debate surrounding the distinction between 'weak' and 'strong' AI is not a subject of discussion in this paper. The distinction between the two is immaterial to the present technical-scientific examination of artificial intelligence.

---

4   Gabler Wirtschaftslexikon.

5   VDI-Statusreport Künstliche Intelligenz, Oktober 2018.

6   Stuart Russell, Peter Norvig: Künstliche Intelligenz: Ein moderner Ansatz. München, 2012. Online-Ressource: http://aima.cs.berkeley.edu/.

7   Glossary of the Plattform Lernende Systeme; https://www.plattform-lernende-systeme.de/glossar.html.

8   IEC White Paper: Artificial intelligence across industries; 2018.

AI from the point of view of Industrie 4.0

From an industrial point of view, AI technologies are to be understood as methods and procedures that enable technical systems to perceive their environments, process what they have perceived, solve problems independently, find new kinds of solutions, make decisions, and especially to learn from experience in order to be better able to solve and handle tasks.

In terms of its function, AI is understood here as a technological form of the human ability to make decisions. It is not intended to copy human behaviour. AI technologies are meant to be used to increase the efficiency and effectiveness of industrial processes. The primary goals of artificial intelligence are to reduce costs, save time, improve quality, and enhance the robustness of industrial processes. But at the same time, artificial intelligence also allows the revamping of production processes and their adjoining processes from the ground up, the enriching of one's own products or services through or with AI, and the implementing of novel business models. These goals will be easier to reach the better industrial processes are equipped with adaptation and problem-solving abilities. The degree of autonomous action in industrial processes therefore depends on how independently a system can automatically master complex situations within the bounds of specified system boundaries. As a rule, the mastering of complex and complicated processes requires knowledge gained through experience and intelligent procedures. Therefore in addition to the simple if/then routines and the classic automation and control procedures, AI appears to be especially suitable for the mastering of complex situations in industrial processes.

In a study sponsored by the Federal Ministry for Economic Affairs and Energy in conjunction with the PAiCE technology programme, businesses in the manufacturing industry claimed that around one-third of the value-added is expected to be linked to AI applications, for example through the attainment of higher degrees of autonomy in the areas of purchasing, production, engineering, distribution, or (after sale) services.[9] But despite such expectations, the actual intensity of use of AI in industrial enterprises is still quite low. The reason for this lies in the enormous changes and expenditures needed to integrate AI applications into corporate structures and along the entire value-added chains. But where it is being used in large corporations, AI applications tend to be found in the areas of robotics or resource management. In SMEs, it is found in the areas of knowledge management, quality control, or optimisation of purchasing supply chains. These applications are based on computer vision, natural language processing, and action planning and optimisation. The actual value of AI applications that are realisable through AI technologies is in the attainment of systems with higher degrees of autonomy and the value-added associated with this.

This current reticence in deploying AI technologies coexists with an expectation that higher levels of autonomy are ultimately going to make it possible to attain goals or functions that are not yet attainable by human beings today. But it is not possible at the present time to make any kind of detailed estimation of all the implications of AI usage. Apart from any possible efficiency gains in relation to the controlling and optimising of production facilities, the biggest impact that AI has is on the organisation of employment in the business world. The tendency to use AI applications to replace repetitive tasks or strongly process-related tasks comes with the creation of new kinds of job profiles and job fields. New forms of cooperation among human beings and between human beings and machines are also being created. The structure of employment on the whole may therefore change in the long run, for example once simple forms of routine work are performed by AI and when AI becomes better, more reliable, and more cost effective than human manpower. But there are already many areas today where artificial intelligence could create some free space for human beings. Examples include those areas where people are reaching their limits due to permanently heavy workloads or excessively high levels of complexity.

---

9   'Potenziale der künstlichen Intelligenz im produzierenden Gewerbe in Deutschland', a study of the Federal Ministry for Economic Affairs and Energy. Carried out by PAiCE supporting research of the VDI/VDE Innovation + Technik GmbH, 2018.

Technological development of AI

The development of KI first began in the second half of the 20th century. It began with the phase of hand-coded AI, the goal of which was the accumulation of knowledge. The focus at that time was on the development of expert systems based on logic-oriented or functional program languages, such as PROLOG and LISP. These knowledge-based approaches evolved into large ontological projects, which were strongly supported right into the first decade of this century, especially in certain areas of application. But with the availability of large volumes of data, large computing capacities, and modern algorithms, machine learning has created new standards in the last ten years and has forced knowledge-based approaches into the background. What the future model appears to be at the present time is one that combines data-based and knowledge-based approaches. The massive quantities of data that are needed for statistical learning alone are in some sectors simply not available or they obstruct the view of interesting fringe cases responsible for unforeseen forms of behaviour in the field. Knowledge-based approaches are also necessary for the specification and the verification and validation of machine-learning applications.

The AI technologies currently being used are divided into rational or behaviour-oriented models.[10] Where humans are communicating with machines, behaviour-oriented AI technologies are generally used, for example natural language processing for machine translation. But in industrial processes involving such things as cost optimisation, object or behaviour recognition in a particular environment, or complex planning processes, the tendency is to develop rational AI technologies. This includes computer vision, action planning, and optimisation.

The driving force today behind AI applications, and one of the main focuses of AI research throughout the world, is machine learning (ML). Machine learning involves the statistical learning of the parameterisation of algorithms for highly complex cases of use. By means of machine learning, the system – using learning data put in beforehand – recognises patterns and principles with respect to the compiled process data. With the help of appropriate algorithms, ML

can be used to independently find solutions to problems that arise.

ML is divided into three fields – supervised learning, unsupervised learning, and reinforced learning. Its more specific (partial) applications are regression and classification, structure recognition and structure prediction, data generation (sampling), and autonomous action.

In the case of **supervised learning**, the system is trained through the correlation of known input data and its corresponding known output data. Of central importance here is the availability of correct data, because if the system is trained using flawed examples, it will learn flawed correlations. In the case of **unsupervised learning**, the system is also trained through using example data, but with input data only and without a correlation to known outputs. It learns how to form and extend data groups, what is typical and where deviations occur. This allows for the defining of use cases and the detection of errors. With **reinforced learning**, the system learns through the process of trial and error by recommending solutions to given problems and receiving either positive or negative feedback to such recommendations. Depending on the reward mechanism involved, the AI system learns to perform appropriate functions.[11]

The learning of hierarchical structures of characteristics in successively higher, hidden network layers is referred to as deep learning. The analysing of complex data quantities is the most significant field of application of supervised and unsupervised learning.[12]

The next phase will centre around the mastering of context-dependent tasks using the appropriate AI. Whether and when this next level will actually be reached is an open question. To do this, machines would have to possess knowledge about the world, i.e. a kind of 'world model', or they would have to have a kind of artificial morality engendered on

---

10   'Potenziale der künstlichen Intelligenz im produzierenden Gewerbe in Deutschland', a study of the Federal Ministry for Economic Affairs and Energy. Carried out by PAiCE supporting research of the VDI/VDE Innovation + Technik GmbH, 2018.

11   Study conducted by the VDMA (Mechanical Engineering Industry Association of Germany) 'Machine Learning in Mechanical and Plan Engineering'.

12   https://www.mckinsey.de/news/presse/kunstliche-intelligenz-potenzial-von-neuronalen-netzen-gigantisch.

their own. At the present time it is unclear as to how this could come about or be created. AI applications are also often perceived as 'black boxes', as a kind of machinery whose functionality in retrospect is not completely decipherable by human beings. This raises questions as to how artificial intelligence can be used in a controlled and comprehensible manner. In other words, and in technical language, the question is how such systems can be specified in advance and verified and validated after the fact. This is particularly important in the case of industrial processes whose comprehensibility (e.g. in cases concerning applications of functional safety) is a significant building block in engineering and operational processes, in qualification and acceptance (of performance) processes, and in standardisation processes.

Why AI is interesting

Artificial intelligence is often something considered obscure or even mystical by non-experts. This is on account of its being ascribed with attributes that compete with human skills. It generally functions in ways that are not clear to us, and the comparing of it to human intelligence triggers bewilderment in us. We are usually very impressed though by the ability of AI to process huge volumes of alternative options with incredible speed, for example in the case of the game 'Go'. But all of the known applications of AI today are highly specialised and have no general applicability. Every new task, prior to the actual application, requires extensive training, which might continue into the usage phase and through which the AI system is continuously optimised. These applications, which cannot be carried out by humans or cannot be carried out fast enough by them due to the amounts of data involved, are usually applications that operate within the bounds of pre-defined rules and system boundaries. This is why a kind of 'savant syndrome' is often attributed to AI. The Go-playing AI cannot steer a car, and the AI that can recognise images cannot understand spoken language. Therefore the direct replicating of human thinking by computers is not a current topic.

The high speed at which today's AI systems can process data and come up with solutions is far beyond what human beings are capable of. Although human beings are universally more intelligent than machines, they are limited by the time it takes them to do things and the possibilities they have of perceiving data through their sensory organs.

For example, human beings have no problem recognising patterns or errors. But when it comes to large quantities of data, e.g. large amounts of rapidly changing process signals, they are unable to cope. The same is true for continuous (monotone) tasks, such as checking outer packaging (for defects), in which cases the human observer quickly suffers fatigue.

AI can assist the automated tasks in industrial processes and the interactions between human beings or between the ever-increasing agility of the work of teams of human beings, and can even create efficiency. This also requires sufficiently powerful and efficient sensors for making the relevant data needed available, a defining and interpretability of the data for compiling the information, and flexible communication infrastructures for transmitting the data to the corresponding edge and backbone systems. This serves to make processes more efficient, and it makes use of the capacity of transparency for the benefit of all participants.

AI is added on to this capacity. The available data is used for matters of problem-solving, exposing of unforeseen occurrences, or for simplifying correlations that are too complex for human beings as such or too complex for their current qualifications. Such matters can arise through the tasks themselves or through the challenging nature of some projects. Projects often face the problem of having to master complexity and, in the planning, have to ascertain costs, deadlines, and resources and have to comply with these.

New opportunities created by AI

AI technologies are said to have a great deal of potential. They can improve the quality of industrial production, reduce costs while simultaneously reducing production times, and increase the robustness of the work processes. Users are also anticipating sustainable potential for themselves through cutting back on resources, optimising of energy consumption, and improving the coordination of logistic processes. At the same time, artificial intelligence also allows the revamping of ideas, product-creation, engineering, production processes and their adjoining processes from the ground up, the enriching of one's own products or services through or with AI, and the implementing of novel business models. AI therefore has implications for all industrial areas and thus for Industrie 4.0 as well.

The efficiency and use of AI thereby depends on a number of specific factors. Of central importance is the quality of the information and the availability and interpretability of it in the learning phase. Equally important are the professional qualifications of the people that develop AI procedures, that transfer them to applications, or that evaluate recognised abnormalities or dependencies or prove and combine them with evaluation schemes.

A major portion of industrial production is already making extensive use of simple forms of automation for certain

tasks and in certain areas, e.g. for processes that can be statically programmed and that are always executed uniformly. In the future, increasing demands for product diversity, for process flexibility, or on costs will create increasing demands for more flexible and independent automation. In its final report, the forum 'Autonome Systeme' (advisory committee of the Hightech-Forum) therefore provided a detailed presentation of the significance of autonomous systems for the economy, science, and for society.[13] The bases for the implementation of autonomous systems are sensorics, robotics, and machine learning. For use in industrial production, it is of particular importance that the autonomous system be capable, in the sphere of human-machine interactions, of adapting to the behaviour of the human beings working with it. Several autonomous (sub-) systems can together form one overall autonomous system of a higher order.

'Autonomous systems' are defined in the technical literature as systems that are capable – without any manual intervention – of adapting their behaviour to unforeseen events during operation.[14] From the point of view of industry, the use of AI technologies should reduce the necessity for human intervention in industrial processes and in the use of intelligent technical systems, and therefore should increase the degree of independence within the defined system boundaries. And in the case of ever-increasing complexities of technical systems, such technologies should make it possible for a human being to make a decision in the first place.

Intelligent technical systems are characterised by their ability to connect with each other, to coordinate themselves agilely, to continuously modify themselves during their life cycles, and to demonstrate robustness in the face of outside influences. During their use, they use and generate a diversity of information and become increasingly more complex, in global value-added chains and when enriched via customer-oriented services and abilities.

The use of AI can enable a system to fulfil tasks independently and without human help within clearly specified system boundaries. The value of AI technologies is most evident when performing unforeseen tasks that demand strong adaptation and problem-solving capabilities. This is not the case when it comes to automation programmed using if/then rules to execute certain actions. Although this does enable autonomous action, it is limited because it can only react to predicted events due to the absence of the decision-making capacity in complex, intelligent technical systems. The same applies in many cases to programming using classic automation and control engineering.

Autonomous systems are found all along the value-added chain and at its preparatory stages: Beginning with requirements analysis right through to engineering, purchasing, supply chain management, virtual commencement of operations, production, marketing, use of products and production systems, and recycling. This applies to the human/machine interaction and the machine/machine interaction, as well as to logistics and process chains. Autonomous systems have clear value propositions to offer to industrial production: The primary reason for increased effectiveness and efficiency in industrial processes and applications is the autonomy of the processes. In the case of the manufacturing industry, the value propositions for the end customers are not found in the AI itself; they are found in the production facilities modified through AI during the entire life cycle of such facilities.

---

13   Forum Autonome Systeme: Chancen für Wirtschaft, Wissenschaft und Gesellschaft, Abschlussbericht 2017.
14   D. P. Watson, D. H. Scheidt, 'Autonomous Systems', Johns Hopkins APL Technical Digest, Vol. 26, No. 4, 2005.

# AI-influenced levels of autonomy in industrial production

The following description of levels of autonomy in industrial production is closely connected to the topic of automation. Large sections of industrial production are already equipped with automation systems, from smaller sub-areas of production right through to entire production facilities. In order to describe this transition from classic programming to autonomous systems, it is important to establish an appropriate taxonomy. This enables the providers of automation technology and their customers to define where they are now and where they would like to be in the short-to long-terms and everywhere in between. This taxonomy is primarily based on two dimensions: the scope of the automated tasks (number, complexity, duration, etc.) and the role played by human beings in this constellation. The following classification of industrial processes is, in terms of autonomous character, based on these two dimensions. It is still a very rough classification, but it is nevertheless capable of being refined and concretised in further steps for a variety of other industrial domains.

The strength of the application of AI technologies is manifested in the intensity of the autonomous industrial processes used, which is linked to the ability to adapt and solve problems. Autonomous action by means of AI technologies is the major force behind increased effectiveness and efficiency in industrial processes and applications. The degree of autonomous action therefore depends on how independently the system can master complex situations – automatically and through self-learning.

AI clothes a system with the ability to adapt and solve problems within defined system boundaries, which is accomplished through training and learning processes. The application of AI technologies then reveals itself in the intensity of the autonomy of the process steps in industrial processes. The level of autonomy correlates to the degree of autonomous action (in machines or processes), which is stronger or weaker depending on the particular application scenarios. The scope of the autonomous action depends on the complexity of the situation to be mastered and the role of human beings in this situation. The following illustration is a simplified depiction of the correlation between autonomy and AI. Levels of autonomy can describe the general state of a system or can be used to describe the desired state of a system in the future. In order to attain a specific level of autonomy, intelligence is needed to develop the system further. Since intelligence is based on knowledge gained through experience, AI is well-suited to providing such capabilities. AI is therefore a technological means of attaining a certain level of autonomy.

Industrial production can be a highly complex affair. It can involve everything right up to a full description of the entire life cycle of a production facility. Although the main factor is the operating of the production facilities, i.e. the controlling of the processes, there are also the areas of planning, engineering, putting into operation, operations, servicing, and shut-downs. These areas are often closely interwoven and mutually impact each other. It is therefore

Illustration 1: General correlation between autonomy and AI.
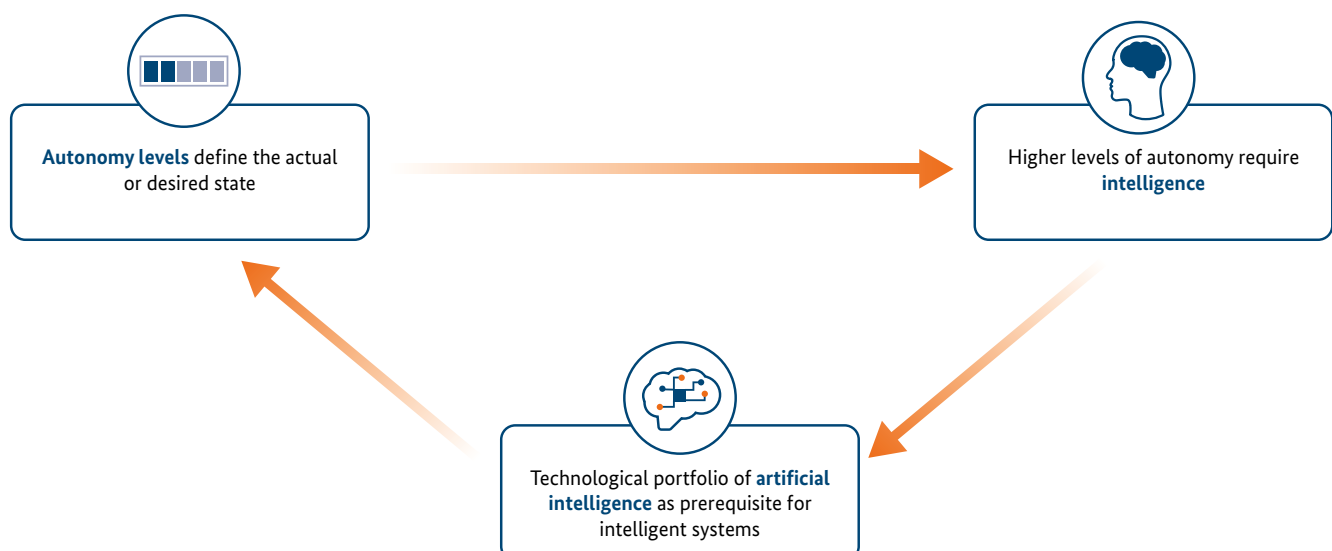
essential to describe autonomous actions on the basis of a graduated model of autonomy, because every industrial system cannot or does not have to attain the same level of autonomy. A graduated model also enables a distinction to be drawn between the sub-areas of industrial production, e.g. process control, process planning, field surveillance, or maintenance. Here as well, different sub-areas can attain different levels of autonomy. The control centre operations may, for example, demonstrate a high level of autonomy, whereas the autonomy of the field surveillance may be quite limited. Also, human beings possess capabilities that are not easily implemented through autonomous processes or that are not replicable for technical, subject-matter, or economic reasons.

Autonomous systems take into account decision-making aspects as well as those involving the execution of actions. With respect to decision-making, the Bitkom committee Artificial Intelligence has developed a taxonomy of thinking, and on the basis of this has defined a graduated, six-level model of automated decision-making – in line with the classification applicable to autonomous driving.[15] This graduated model describes the interaction between human beings and tools that assist or work independently. The project group Artificial Intelligence of the Plattform Industrie 4.0 uses this graduated model. It presents below the taxonomy of system autonomy based on AI, it illustrates it on the basis of industrial processes, and infers from it in-depth knowledge and actions needed for a variety of applications.

Fundamental to the defining of the levels of autonomy is the role of human beings in the value-added process in relation to the scope of the automated tasks. Autonomy Level 0 defines an industrial production operation without AI-based automation pursuant to which human beings maintain full control and responsibility (although an extensive and simple form of automation is allowed to be present). Level 5 on the other hand defines a fully automated production operation controlled by AI pursuant to which the entire decision-making and execution roles are assumed by an AI system. Levels 1 to 4 define the different levels between these. It is important to bear in mind here that the defining of these levels is always made pursuant to

defined system boundaries and with known forms of AI technologies. An execution of an action by means of stronger forms of AI technologies and with undefined system boundaries is not a part of this discussion. There is no way with today's knowledge of predicting whether such systems will ever be used in industrial production.
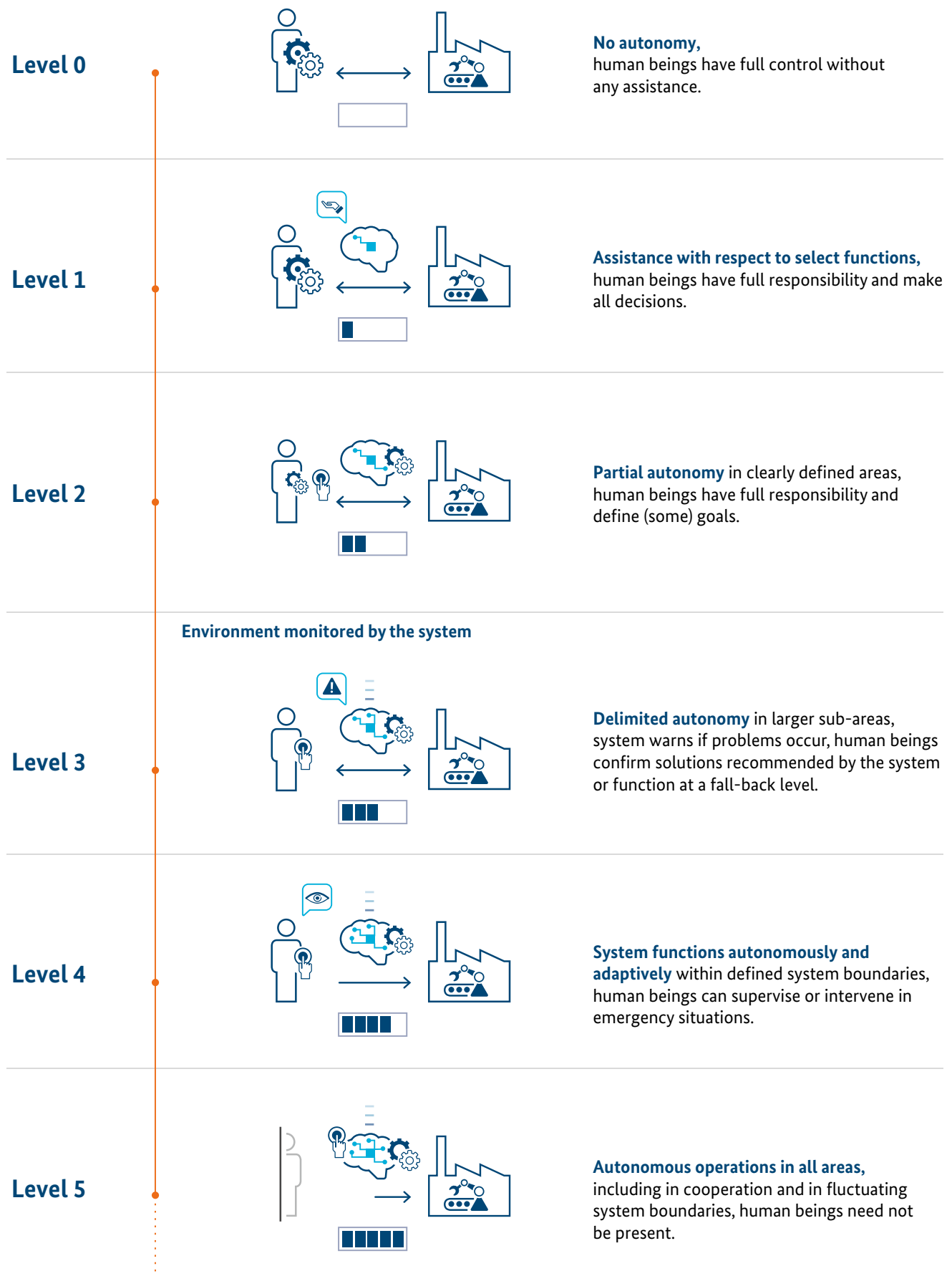
The following illustration summarises the generic definitions of each of the autonomy levels for industrial production. These are understood as the automation of processes in specified system boundaries. The different levels are attained through increased automation. Examples of automation include automatic translation, automated order-controlled production (bicycle handlebars[16]), chatbots (automation of service processes), engineering support, robotic process automation (RPA), and the automatically controlled classic production line albeit for which no AI is needed. The levels of autonomy presented below are purely in relation to the effects of AI as another form of automation technology. The overall level of AI-based autonomy of an industrial process can range from no autonomy at all (Level 0: no AI used in the programming), to partial autonomy (Levels 1 to 4: AI programming is used; human intervention is still needed in the process), to full autonomy (Level 5: AI programming replaces every form of human intervention).

These defined levels of autonomy show a gradual transition of the responsibility for the production operations from human beings to autonomous systems. At Levels 0 to 2, certain (partial) autonomous actions are possible. These are limited, however, and human beings maintain the active control and central responsibility at all times. At Levels 4 to 5, the system assumes responsibility – first for partial areas and partial aspects only and then gradually for the entire facilities. The role played by human beings is largely passive. Level 3 defines a transitional situation in which certain decisions made by the system must still be confirmed by human beings. Levels 3 to 5 place high demands on the reliability of the system, which is why one important fringe condition must be met for the transition from Level 2 to Level 3: The system must monitor its work environment in order to be able to respond to unforeseen events, for example.

---

15   Bitkom guidelines „Künstliche Intelligenz verstehen als Automation des Entscheidens", 2017.

16   Plattform Industrie 4.0: Anwendungsszenario trifft Praxis; Auftragsgesteuerte Produktion eines individuellen Fahrradlenkers, https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/anwendungsszenari0-trifft-praxis.html.

Illustration 2: Generic definitions of AI-impacted levels of autonomy in industrial production.

Level 0

No autonomy,
human beings have full control without any assistance.

Level 1

Assistance with respect to select functions,
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.

Environment monitored by the system

Level 3

Delimited autonomy in larger sub-areas,
system warns if problems occur, human beings confirm solutions recommended by the system or function at a fall-back level.

Level 4

System functions autonomously and adaptively within defined system boundaries, human beings can supervise or intervene in emergency situations.

Level 5

Autonomous operations in all areas,
including in cooperation and in fluctuating system boundaries, human beings need not be present.

The illustration of these classification levels will be demonstrated using the example of a robot that handles parts. Such robots pick up parts and set them down somewhere else, for example on an assembly line.

Level 0

At Level 0, the controlling of industrial processes takes place purely through programmed if/then routines or through classic automation and controlling engineering. This is generally done through a programmable logic controller (PLC). No AI algorithms are involved. The rules are clearly determined by human beings, who maintain full control over the entire process at all times. An extensive set of programmed rules may nevertheless exist, for example machine controls or safety-relevant functions for warranting the safe operation of the facility. But the responsibility for keeping the operation going and for responding to unforeseen demands through the use of intelligence is solely in the hands of human beings, e.g. by creating the rules or by controlling the process.

Autonomy Level 0
No autonomy, human beings have full control without any assistance.
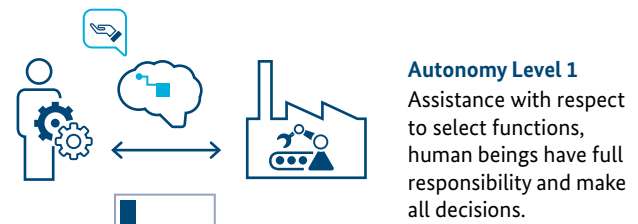
Example: Robot for handling parts
A robot picks up parts from predetermined positions and puts them down at exactly defined positions. At autonomy Level 0, the robot is programmed or designed in a fixed (i.e. pre-set) manner by human beings in the classic way (teach in). Prioritising and selecting the rules are done by human intelligence. The robot works within fixed boundaries (system boundaries).

Level 1

At Level 1, the AI system performs supportive functions and therefore operates as an assistance system. The AI procedures facilitate the interpretation of complex and ambiguous information, for example from human language or images. AI-supported interfaces can provide users with help, protect them from operating errors, and provide them with optimisation options and predictions on the basis of recognised use requests. Assistants are able to suggest meaningful input data and predict the consequences of the input. AI-based assistance systems can also help optimise the work within a team and between teams. The AI allows for recognition of the situation, taking into account the various goals of the individuals involved. It can then intervene in both a coordinating and an intermediating manner. It is also easier to preprocess existing information and to share it within the group.

The goal of assistance systems is to reduce complexity and to guide human beings in difficult situations. In relation to the process as a whole, human beings at Level 1 still make all the decisions and maintain full responsibility for all processes.
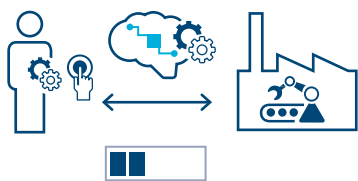
Autonomy Level 1
Assistance with respect to select functions, human beings have full responsibility and make all decisions.

Example: Robot for handling parts
As with Level 0, the robot at autonomy Level 1 is programmed or designed in a fixed (i.e. pre-set) manner by human beings in the classic way and works within the same fixed boundaries. The rules and the selecting of the rules are created by human intelligence, as is the case at Level 0. In contrast to Level 0, the robot's assistance system, which was programmed using AI, recommends goal-oriented improvements to the human being, e.g. process optimisations regarding energy or time. The person then decides the extent to which he/she wants to use the recommended improvements.

Level 2

The idea underlying Level 2 is the automation of simple tasks or the delegation of simple tasks to automation systems. More complex tasks remain in the hands of the operator. In defined areas, to the extent desired, and for a limited period of time, the system assumes control under the supervision of human beings, who also control the results.

The operator supports the system with experiential knowledge, sets the goals, and under certain circumstances defines the intentions behind their actions. The responsibility for the process as a whole is still completely in the hands of human beings.

Autonomy Level 2
Partial autonomy in clearly defined areas, human beings have full responsibility and define (some) goals.
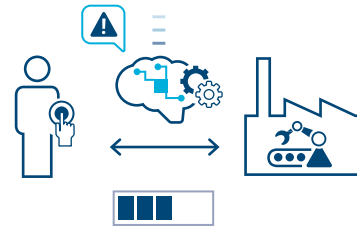
Example: Robot for handling parts
The robot at autonomy Level 2 is still predominantly programmed or designed in a fixed (i.e. pre-set) manner by human beings in the classic manner. In contrast to Level 1, the AI programming allows the system to improve itself within the specified system boundaries and goals. For example, the robot can recognise and pick up parts that are not in the exact predetermined position, or it can independently adjust the frequency rate of its actions. Human beings have decision-making powers in all areas and can intervene if necessary.

Level 3

At Level 3, the operator defines the system boundaries within which the AI-controlled system is allowed to control the automated operations. The system independently monitors its environment within the specified system boundaries. Through the monitoring of the production process in real time, defects and bottlenecks, for example, can be detected. If the AI recognises a pattern from the learning phase, it is able to semi-autonomously execute an appropriate, learned action. An example of this is when it independently activates an emergency stop or puts the facility into secure mode if an acute overload is detected in the sensor data.

Such semi-autonomy not only serves to optimise the production flow but also reduces down time, as problems and defects are often detected in advance or the persons involved are informed about them in real time. This allows for the economising of materials, resources, and replacement of parts. The role of human beings is merely to con-

firm certain recommended solution strategies or to provide support to specific problems, i.e. they intervene when they are alerted by the system.
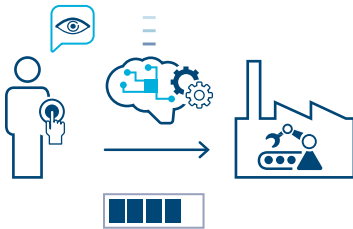

Autonomy Level 3
Delimited autonomy in larger sub-areas, system warns if problems occur, human beings confirm solutions recommended by the system or function at a fall-back level.

Example: Robot for handling parts
At autonomy Level 3, the robot is only partly programmed in a fixed (i.e. pre-set) manner by human beings and is designed to a large extent. In addition to adjusting its own actions, the robot at Level 3 can itself make and implement plans within the system boundaries (e.g. autonomous continuous path control), and can do this in cooperation with the environment, such as with other robot units. It can for example modify the speed of an assembly line. To do this, the robotic system is equipped with sensors for registering its environment. It is able to perceive the environmental context, to adjust movements, and to learn skills. Human beings observe the system's decisions, help with unforeseen problems, and can intervene in emergency cases.

Level 4

At Level 4, the system operates as an adaptive, autonomous system in larger sub-areas (e.g. operating the control centre). It works autonomously and adaptively within known system boundaries and can optimise itself through ongoing learning phases and defined (partial) goals. This allows for improved predictions and problem-solving. A self-optimisation of the production strategy on the basis of specified key data is enabled within the system boundaries defined by the algorithms. Human beings relinquish to the system control of a desired part of the system. Humans have at most a monitoring function and intervene in emergency situations. When the human being fails to intervene, the system will itself take care of a variety of situations according to its own ideas of what is 'right'.
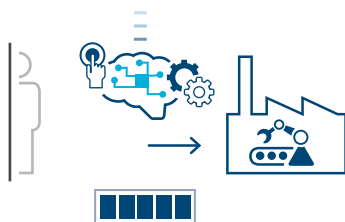
Autonomy Level 4
System functions autonomously and adaptively within defined system boundaries, human beings can supervise or intervene in emergency situations.

Example: Robot for handling parts

In contrast to Level 3, the robot at autonomy Level 4 operates autonomously. To achieve this, it is equipped with all of the sensors needed for registering the environment. It fully perceives the environmental context and reacts autonomously within the system boundaries specified by human beings. Human beings monitor this and can intervene in emergency cases.

Level 5

At this highest level of autonomy, a facility or an extensive process operates in all areas completely autonomously. The system works out self-organised, adaptive solutions and does so cooperatively and within fluctuating but still specified system boundaries. Human beings can be completely absent here; no user interaction with the facility is needed.

Autonomy Level 5
Autonomous operations in all areas, including in cooperation and in fluctuating system boundaries, human beings need not be present.

Example: Robot for handling parts

At autonomy Level 5, the robot works completely autonomously within the system boundaries specified by human beings, including in cooperation with other autonomous systems. For example, with the help of plan-based procedures, the entire process as well as the required sub-processes can be planned and exchanged with the other processes involved or with other autonomous systems. In the case of fluctuating working parameters, the plan is dynam-

ically revised and communicated to the other machines. Human beings need not be present. In emergency cases, the system independently puts itself in secure mode.

Example: Quality assurance in the manufacturing of rotary blades for wind-powered installations

The example of quality assurance in the manufacturing of rotary blades for wind-powered installations clearly brings to light just how beneficial AI support, using machine learning, is when inspecting rotary blades directly following production. In order to avoid liability for manufacturing defects, extensive quality inspections must be carried out. In the long rotary blades made of glass fibres, no cracks, contortions, material inhomogeneities, or other defects are allowed as these will lead to fatigue failures. These do not show up as cracks as in the case of metals, but as crevasses extending across a large surface and are responsible for ruining whole rotary blades. Therefore to prevent this, an optical scanning of the surface of the blades is performed. The inspection formerly carried out by human beings took a lot of time, and defects were often overlooked due to fatigue on the part of the inspectors. Such a purely human inspection is classified as Level 0 and involves no AI.

In order to improve the performance of the inspection and the quality of the execution, the inspection is now performed using an optical scanning process with an AI-supported evaluation of the images.[17] In the graduated model set out above, this quality inspection without any further AI support is still classified as Level 0. This system can be optimised through an appropriate ML training phase and defects can be completely eliminated in approximately 80 % of the surface. The remainder is inspected by human inspectors. Improvements have been identified in terms of speed, costs, manufacturing capacity, and even with respect to employee satisfaction. The inspection system at Level 1 assists the human operators by providing them with the complete results of the scanned images and by pointing out any anomalies in the images. The human observer then marks the defective areas for the further processing.

At Level 2, the AI-assisted inspection system itself makes a list of the defective areas, adjusts its scanning speed to the surface conditions, and learns during implementation. New images of defects are added by AI experts to the neural network in order to continually improve precision and to ulti-

mately reduce 'false positives' to zero. The need for human intervention is therefore continually reduced, which allows the workers to devote more of their time to the tasks they are actually qualified for.

The inspection system described is already in operation. Such a quality assurance system can also be brought to higher levels in the future, for example by combining it with another system that automatically repairs the defects discovered. The optical system would independently communicate the defective positions to a repair robot, even perhaps with information about the nature of the defect (how widely spread, how deep, etc.) and with a recommendation on how it should be repaired. At Level 3, the known repair procedures could be executed by human experts. Through additional learning phases, including during operating times, the overall system could be further developed into an autonomous manufacture-inspection-repair system. A system that would be capable of detecting defects while the glass fibres were being laid and of correcting them at the time of production already. Such systems would then be allocated to Level 4.

Example: Increasing operability of machines and production facilities
This example illustrates how AI can be used in connection with the area of what is often referred to as 'predictive maintenance'. The goal here is to design the operating, the maintenance, and the eliminating of defects of machines and facilities as efficiently as possible. Maintenance is understood here as the planned work needed to maintain the operability of machines and production facilities. AI can be used to assist with issues concerning monitoring, explanations when incidents occur, and the making of predictions. It could also make recommendations regarding measures to be taken and could ultimately act on them.

The fixed (i.e. pre-set) programmed machine code for the automating of machines and production facilities has always contained a large share of evaluations for monitoring and diagnosing process steps, operating resources, and quality specifications for the manufactured products. If thresholds are exceeded, such systems respond in the ways defined by human beings, e.g. by setting off an alarm, sending an e-mail, or other such measures. The causality and 'intelligence' of the rules are defined and designed by machine experts. This is allocated to Level 0 and has no connection to AI.

Benefits could be obtained through the use of AI resources in the area of machine operation, benefits that go beyond the fixed (i.e. pre-set) programmed automation and beyond the resources available to human beings. An example of this is the anomaly detection for machine data. This involves monitoring complex data from machines (large number of various sensor values and large amounts of data) using artificial intelligence and comparing it to learned data patterns illustrated in algorithms. The benefits are:

- The 'training' of the algorithms is done empirically. This means that no expert knowledge is needed for the causality of the machine processes.

- The starting point for training the algorithms is the sensor data of the machines in the various 'normal' states of the production facility. Therefore no comparison between known errors and existing error sensor-data is needed – which are often not (yet) available.

- Because anomalies are usually identifiable much earlier in the machine data than in the physical machine itself, an earlier response to them is possible.

- In contrast to 'simple' rules, AI is also capable of analysing highly complex data. This includes such things as reviewing data correlations from thousands of sensor values, for example.

- Machine experts can also be offered decision-making support through the identification of the machine components concerned and with visualisations of the causal sensor values.

- The systems continuously gather the decisions and thereby expand their knowledge.

- Maintenance processes often focus on individual machines and parts of production facilities. This can arise from the combining of machines of different manufacturers or from the complexity of the requirements. With inter-machine AI, value chains can be developed and observed as a whole.

In addition to the information needed to operate the machines, the operators are offered further information about the availability of operating resources and the quality of the products produced. At Level 1, the information

needed by the machine operator is extracted from the wealth of sensor data or process information. The assessment of and responses to it are made by human beings. Should a malfunction occur, the cause of it is identified, information is given on it, and attention drawn to the instructions on how to remedy it.

At Level 2, measures are activated in defined areas autonomously by human beings. This can be used to make the machine itself operate more robustly, e.g. by making adjustments in relation to the wear and tear of tools, or to assist the operator of the machine. Should a malfunction occur, machine operators can be provided with remedial recommendations that they can implement themselves. In the case of costly tasks, an alarm could automatically trigger the commissioning of a specific service technician. In such cases, plans of action for these commissions could be in place, which could take into consideration things like the required expertise of the technician, the tools available,

or the physical proximity of the technician to the place of work. Important information regarding the malfunction and possible recommended solutions could also be handed over to the service technician. Other examples are found in the area of maintenance work. AI can be used here to keep track of operational behaviour in order to adapt maintenance cycles to the real use of the machines and production facilities. In such cases, maintenance work could be carried out more productively. Goals for AI applications include:

- performing maintenance work only when it is actually needed and not at fixed intervals,

- remedying problems before damage occurs or at times when it does not significantly disrupt work flows, e.g. after an order is completed or during work breaks,

- combining of various work on machines and production facilities with the necessary resources at agreed to times.

Human beings define the system boundaries

A wide field for AI in industrial environments is the analysis and interpretation of sensor data, which is distributed throughout the machines and production facilities. It records the status of all different aspects of the machines and performs actions in process work flows on the basis of it. Its central purpose is to identify correlations that are not obvious, e.g. to enable predictive maintenance work. AI is also used in process-, logistics-, and energy-optimisation of industrial processes, for example when complex interrelated mechanical setting parameters have to be adjusted in response to fluctuating conditions in the environment. The 'internet of things', i.e. distributed data-suppliers and data-users capable of communicating with each other, is the basis for AI and places high demands on the monetary benefit in the implementation of business models.

AI includes a portfolio of applications and technologies that enable the realisation of autonomous functions and systems as value-added assets. Autonomy always takes place here within system boundaries defined by human beings. Human beings define which level of autonomy is to be attained for which overall system, and they decide the areas and functions within which AI is allowed to operate. Examples of overall systems include real estate properties, production facilities, or a precisely defined task area. The boundaries of the overall system should not be defined too narrowly, however. Functions include such things as the control centres of a facility, the engineering of a production line, or logistics planning. The level of autonomy of a system is not necessarily determined by the technical limitations of the AI alone. It is also influenced by aspects such as the underlying legal framework, estimation of the advantages and disadvantages of human actions, or the demands of data protection.

Autonomy levels define the system boundaries of the AI technologies used, i.e. which tasks are allowed to be executed. A self-driving car at autonomy Level 3 can for example independently take over the driving on the highway but not the exiting from such highway. From Level 3 upwards, the surveillance of the environment by the AI system is added to the specified, process-related system boundaries. The system must be able to independently perceive its environment and make 'real' autonomous decisions within the specified system boundaries and without further human intervention. This raises important legal issues, for example regarding the liability for and the comprehensibility of responsible decision-making. These are issues that need to be debated by society as a whole, beyond the spheres of industry and politics and regardless of whether the issue is autonomous driving or the use of AI technologies in industrial production processes. The level of autonomy therefore determines, in relation to AI, the general system boundaries in which AI is allowed to operate, which therefore also includes the non-technical boundaries of the AI technologies used, such as legal or ethical boundaries.

Didactics of artificial learning

For simple automated processes in industrial processes, the classic programming and additional human intervention suffices in most cases. But in the case of more complex processes with higher decision-making demands, a programming based on AI technologies is advisable. In contrast to classic programming, however, the use of AI means that there is no expectation of the exact same result being achieved each time. This is because of the extensive data and information pools used for the learning process. What is strived for instead through the use of AI is a continual optimisation of the processes and the making of 'right' decisions when problems arise. Allowed here are also solution recommendations by the AI system that are not those exactly expected. But the decisions made must always be plausible and must as far as possible correspond to the intended courses of action.

The discussion surrounding the impact of AI on industrial processes can be divided into two phases: decision-making (rule-creation/design) and execution of actions (rule-execu-tion). These can be equated with the skills needed for making decisions (determined at the time of programming) and the implementation of the decisions through actions (have an effect at the time of execution). In the case of AI-assisted design, the system is trained through an initial learning process and, together with classically programmed sub-areas, is enabled to execute an industrial process. The AI-impact decisive for the rules concerning 'execution' can be classified using the autonomy levels described above. In the 'designing' phase of the industrial process to be initiated, human beings in the role of 'supervisors' decide in advance which data is allowed for the first learning phase. They must also decide which data the AI system is allowed to receive for a further learning phase during the controlling of a process. Although data is needed for optimising the process, an overfitting must be avoided. This is because an overfitting of the data analysis can have a negative impact on the existing problem-solving model. The degree to which automatic learning is allowed can be roughly divided into three levels:

Illustration 3: Rule cycle for a classic execution and an AI-based execution of an industrial process.
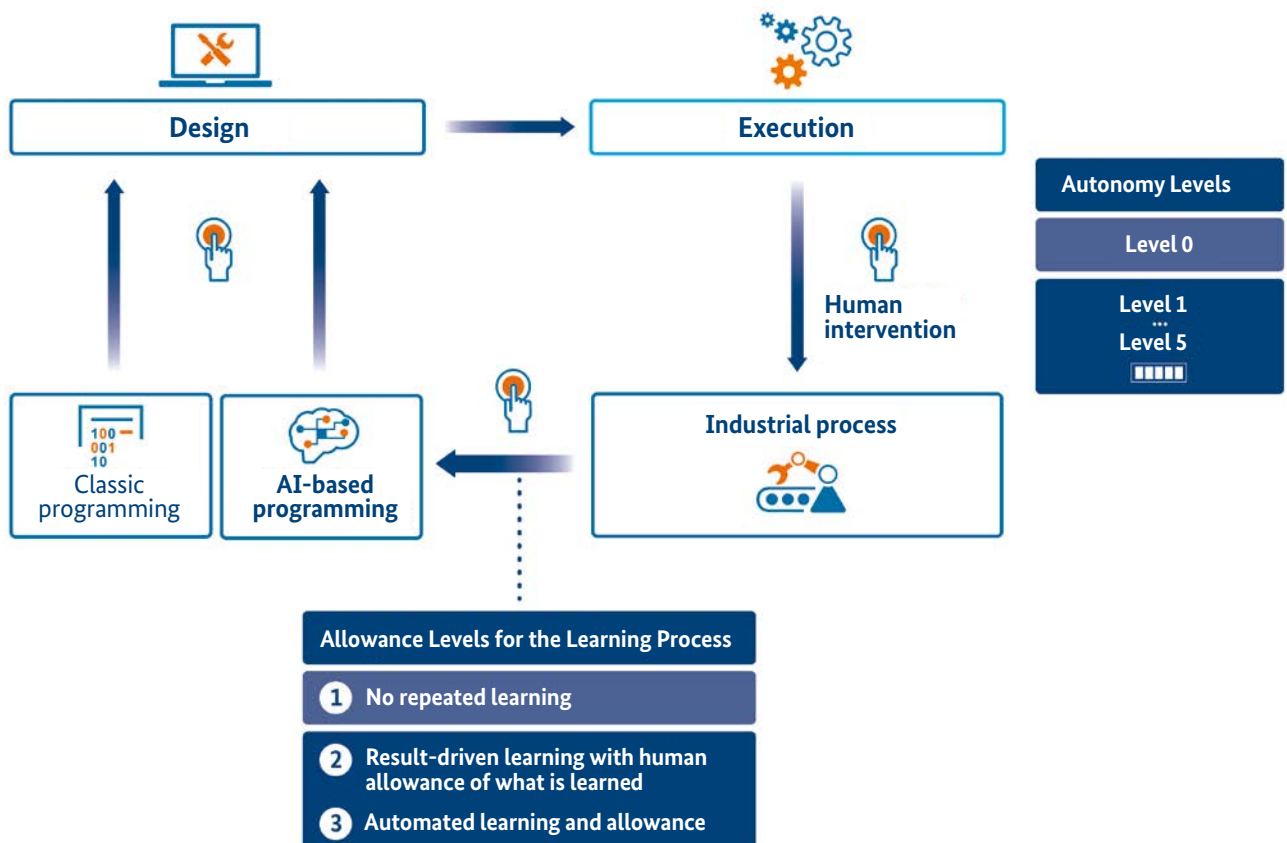
Allowance Level 0 – Manual learning (no further learning allowed)

New learning takes place only if manually extended, e.g. for creating new rules. No automatic extension of knowledge takes place.

Allowance Level 1 – Result-driven learning with allowance by human beings

The AI independently collects additional training data and through continual learning autonomously generates new knowledge. The use, however, of any new knowledge must be manually reviewed and allowed by human beings.

Allowance Level 2 – Autonomous learning (automatic allowance and learning)

The AI independently collects data, generates new knowledge, and uses this knowledge to adjust its behaviour (rule-making) within the specified system boundaries. Rule adaptation takes place completely automatically without any manual intervention but only within the framework of previously specified boundaries.

Today's automated processes in industrial production are largely created on the basis of simple if/then rules and classic automation and control engineering. This requires a thorough penetration of the requirements and the problems by the rule maker. The rules created must cover all of the work processes along the value chain and as many disruptive process-anomalies as possible. But the rule-maker's ability to penetrate the requirements and the problems is limited by the reproducibility of the complexity of such requirements in programming and engineering systems and by the powers of comprehension and levels of experience of the participants. The increasing complexity and dynamism of systems demands more than what human beings are capable of. Programmers may become better and better through more and more experience, but if they are no longer able to penetrate and replicate all of the process flows, then the performance capacities of the technological systems and their interactive possibilities may very well increase, but the potential of the overall system will not be fully unlocked. It is this gap – i.e. between the potential [of the overall systems] on the one side and the tasks to be solved by human beings through programming and engineering on the other – that is being filled by AI.

The functional impact of AI manifests itself in the execution of processes (engineering, machine control, pattern recognition, logistic processes, invoice checking, etc.), but these functions are determined through the creation of a set of rules. A programmer defines the system boundaries, the kind of programming, and the form of human intervention. The programming of the execution can be influenced by AI to larger or lesser degrees. It is completely irrelevant, however, for the execution whether the rules are created through classic programming (if/then scenarios, control regulation, etc.), AI algorithms, or through human intervention. Decisive for the classification in a specific autonomy level is the necessity of human intervention in the overall process. At a higher level of autonomy, execution by human beings decreases and the influence of AI increases. In such cases, comprehensibility, repeatability, reproducibility, and plausibility are possible factors when using AI programming. If, through the system controls, an unintended action occurs, then – depending on the specifications of the allowed system boundaries and on the intended level of autonomy – it must be possible to adjust the controls or even to activate an emergency stop.

What is also needed here is a controllability concept. The human being must be capable of mastering even those tasks referred back by the machine. For example, in cases where the machine is not able to perform a task and refers the task back to a human being for a decision, then it must do this in a timely manner so that the person is able to respond as he/she deems fit.

Qualitative data with context makes AI possible

For AI technologies to be able to support human beings in future industrial processes, the systems must be trained through learning processes. Such learning processes are based on inputs of appropriate data and the use of system-oriented algorithms. More important than data quantity is data quality, especially the particular context of the data (annotated and curated data). Following this, the trained systems – in combination with the appropriate technology – are used to supplement human capacities, both the physical as well as the cognitive and decision-making capacities. It is mainly the cognitive capacities that AI is assisting, in which case it is not only for saving time and money but also for improving quality and for liberating human beings from having to perform simple, repetitive tasks. The idea is that human beings are to have more time for the creative

planning of intelligent processes. They are not supposed to be replaced, but are supposed to use their wealth of ideas for creating new models of technical processes. Artificial intelligence, as a human tool, creates new perspectives for this and can become one of the core building blocks of Industrie 4.0. This requires a mechanical training phase, central to which is the availability of high-grade data and information. The more such context-based training data is available, the better the AI-controlled machines will be.

For the upholding of value propositions in industrial processes, the quality of the data used cannot be stressed enough. This is why the purely data-driven approach to using data pools is simply inadequate when it comes to training AI systems, finding new solutions to complex processes, and generating unique selling propositions for business enterprises. In order to select good data from available sources of general information or from a company's internal or cross-company databases prior to the learning phase, an intensive period of preparatory work must be performed by human beings. Such work necessitates specialist expertise in the classifying, analysing, and defining of qualitative high-grade data. Such data specialists must therefore not only be knowledgeable in their field but must also possess context knowledge in order to train AI systems. For Germany to preserve its status as a competitive technological nation in the future, the creation of educational qualification programmes that combine AI expertise with domain expertise is essential. Sector-specific specialists could be taught to extract specific, qualified data from data pools and then to feed it into AI systems in a context-based manner.

Data availability

For the training of AI systems, it is not large data pools that are needed. The success of such systems depends instead on the availability of pools of context-based, logical, and qualitatively high-grade data. Unlike the consumer sector, the data needed for the manufacturing industry is of a more specific nature. Companies generally have their own individual concepts for their production facilities, and to transfer the machine data of these to the facilities of other companies is not a simple task. Even in the contemplated 'batch 1' (customised) production, there is generally only one qualitatively 'approved' part on a production facility at any one time, which even for the AI-training phase of the facility is of limited use only. For data to become useful information, it has to be defined at least in relation to its function and abilities.

With respect to the use of a suitable data pool or information pool for the conditioning and optimising of a system, numerous hurdles must be overcome. For example, data that is pooled from different companies or from machines of different manufacturers usually comes in a variety of formats. What is needed in such a case is an infrastructure that is capable of compiling such unstructured data and data that is accruing rapidly in real-time. But because companies are generally suspicious of each other and want to guard their know-how, they are cautious when it comes to exchanging data. There are no rights of ownership to data. It is only protected against unauthorised distribution/use. Data protection is therefore in a company's own best interests and is normally regulated in confidentiality agreements. As the law currently stands, data exchanges especially between competitors might amount to infringements under antitrust law. If data or information is regarded as a product, then any use of it in a foreign country must comply with export control laws. And of course all existing local laws in force – such as the provisions on the storage of personal data or the General Data Protection Regulation – also have to be observed. It therefore still needs clarification just how the exchange of data between manufacturers of industrial products and their users in industrial production facilities is to be legitimated, motivated, and economically realised for all stakeholders.

The challenge is therefore to be able to supply data in adequate quality for the development and testing of AI technologies without encroaching on the sovereignty of the possessors of such data. These two sides appear at first to be irreconcilable. In order to do justice to both interests, the need for making data available and the use of data must be clarified within the context of the existing interests. The argument against not giving out data for personal or economic reasons is often based on reasons of intellectual property, use by competitors or by uncontrollable third parties, or on a lack of appreciation of what data is worth. On the other side, the taker side, is the general public interest in obtaining good data for the further development of AI or for the generation of new business models. This taker side favours technological progress and innovation using as much data as possible.

Global examples of the various possible methods for resolving this conflict of interests exist, some of which are on the market. For example, in special economic areas with restrictive regulations (e.g. regional legislation), the publication of data is mandatory. In other areas involving highly personal forms of use (e.g. social media, search engines,

etc.), the providing of data is tolerated, which therefore allows the operators the nearly unrestricted use of it. It might also make sense to enlist the publicly available open data platforms for providing data.

Another possible method would be to make data partially available on a trust basis. The trustee would make data of general usefulness available without identifying the origin of it or without any identification of a person. Such data could be used by the general public for the further development of AI or for the creation of business models. Groups of users or individual consortiums could also agree, to a certain extent, on a set of data-usage rules and could enter into contractual agreements with committed companies and user groups. Such a mechanism based on trust presupposes transparency with respect to the rules and trustworthiness on the part of the organisations and structures. Examples of such trust-based organisations could include cooperatives, government institutions, or trustworthy technological infrastructures. Certain areas of statistics could serve as an analogy here, pursuant to which sales figures are collected using a notarization mechanism and published in an anonymous form.

Data is valuable

The data economy, and the new business models based where every participant benefits (costs, quality, time savings, liquidity, flexible use of employees, etc.), is going to play a major role in the future with respect to the giving and taking of data. General questions are therefore arising regarding how to measure the cost-benefit-ratio of data pooling, what the value of data is in general, and how pricing mechanisms could be developed. Business enterprises should therefore decide in favour of giving their data to a pool on the basis of an appropriate model, because the data-giver benefits from the use of the data pool in the form of added value.

An initial model for the creation of a supervised and high-quality data pool and for determining the value of data could be the institution of the clearing (intermediary) body. Such institutions could function along the lines of cooperatives or on a trust basis with contractually agreed rules and confidentiality principles. Such a clearing body would have to assure that the data was of adequate quality, would have to guarantee a standardised form of data provision, and would have to regulate the monetary realisation of know-how. This would necessitate a complete change of consciousness on the part of business enterprises, i.e. that data is something that should be refined and made use of instead of being possessed.

AI in Plattform Industrie 4.0

Industrie 4.0 covers a wide range of perspectives, each of which is the focus of one of the platform's working groups. The uses and ramifications of AI for Industrie 4.0 are of particular relevance to them and are the topic of the following chapter (and of more in-depth discussions in other papers).

Reference Architectures, Standards and Norms

With the reference architecture Industrie 4.0 (RAMI 4.0), the components of Industrie 4.0, and its administration shell, Working Group 1 – Reference Architectures, Standards and Norms – has set itself the goal of defining the underlying requirements for an Industrie 4.0 system comprised of interacting components. It is based on an underlying interaction mechanism that is built on fundamental standards using interpretable semantics, that allows for a plurality and interoperability of the communicating assets, and that enables the interaction between various manufacturers and users across the different system boundaries. The term 'assets' is understood here to include all operating resources, workpieces, and acting persons.

For key technologies such as AI, which are heavily reliant on large quantities of information from a large number of different sources, the factors listed below are essential. These are made available by Industrie 4.0 and its administration shell:

- Provision of information through trustworthy partners and components in the network. The main idea here is that information about the components is made accessible through Industrie 4.0 network and does not go unused because of potential technical hurdles. Information streams must be structured and controllable. The mechanism must facilitate both a decentralised data-source-structure and centralised data storage.

- For interpreting the information, the context as well as semantics and a grammatical structure for the information are needed. The more explicit the language is, the easier it will be for the participants to communicate with each other. Functions and consequences must be unambiguously defined.

- Accessing, collecting, analysing, and learning from information are the primary uses that AI has to offer to industrial applications. The next important step includes working with the results and refining autonomous processes. This requires the active interaction between data sources, AI learning mechanisms, human decisions, and the implementation of the learning results in controlling or regulating functions. These – as Industrie 4.0 components and therefore also as functions – must be replicated via an administration shell as well. And all of this in systems that are constantly changing.

- Industry has its special demands not only on resources, availability, long-term stability, and robustness but also on construction sizes, environmental conditions, and costs. These may conflict with the requirements of AI, i.e. the need for as much information as possible on a permanent basis.

The basic structure of the administration shells and the interaction patterns resolve some major challenges. AI technologies can make decisive contributions to a system's autonomy and response capacity. The symbolic closing is also of great importance as a key basis for automated decision-making.

The following challenges still need to be tackled:

- AI makes it possible to combine semantics automatically. Languages and semantics change over time and AI mechanisms can enable them to be used more resiliently. Created and explicit knowledge must be constantly validated by domain experts and users.

- Pattern recognition and the detection of anomalies is another key aspect of the use of AI. These capabilities support agile action. Self-learning systems can produce models which expand explicit semantics in a multilayered manner. These are capabilities which create potential for future business models and offer methods which are likely to generate more innovations in efficiency gains and reduced resource inputs.

In overarching terms, the following recommendations and demands to government and business can be derived:

- Business and government must invest sufficient funds in AI research and for testing in test beds.
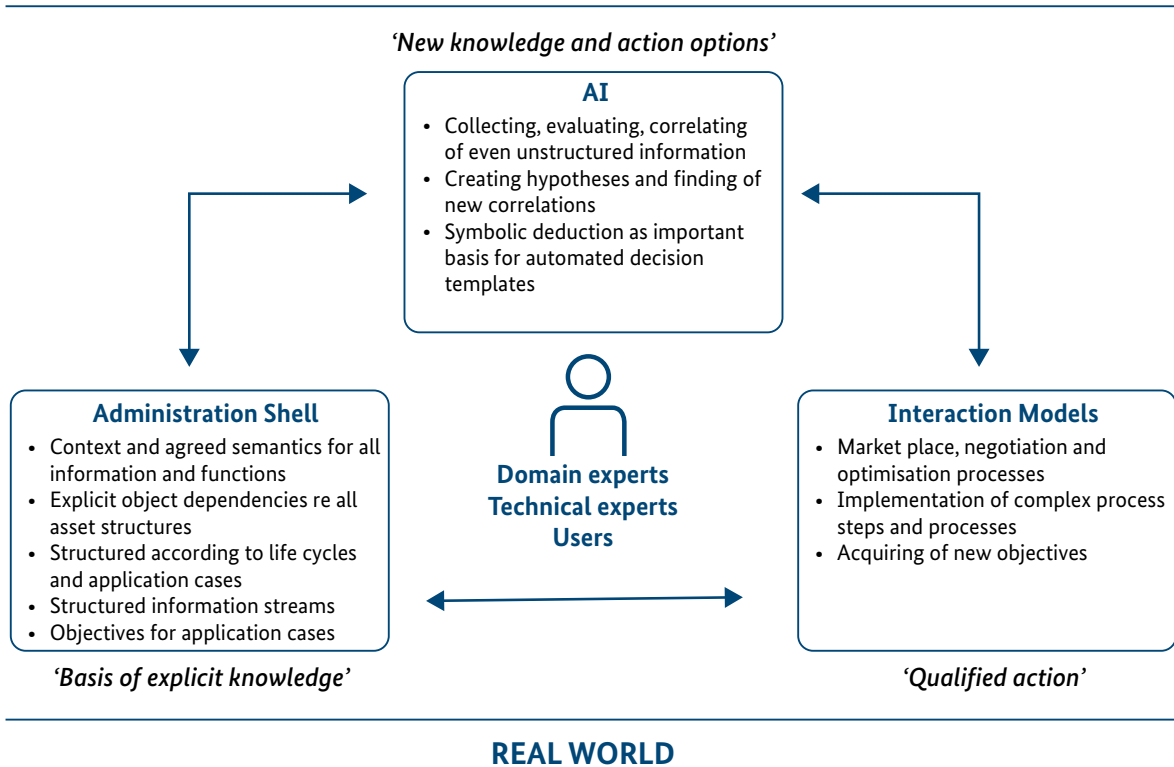
- SMEs need to be made aware of and trained to handle this issue. Training for AI (e.g. data scientists) must be promoted.

- It must be possible to share sample and teaching data between value creation partners (sort of "open data" for AI technologies).

- AI implementations must meet industrial requirements (e.g. computing power).

- AI algorithms must be capable of standardisation so that they can produce validatable results in as many applications as possible.

Technological and Application Scenarios

Working Group 2 – Technological and Application Scenarios – focuses on the impact of existing and forthcoming technologies on their use or on the industrial production with Industrie 4.0. Generic application scenarios were created for demonstration purposes. These scenarios represent an overall framework for the application of Industrie 4.0 in the digitised industry and illustrate impacts and requirements from the perspective of the users and working groups of Plattform Industrie 4.0.[18]

In the further development of this working paper 'Artificial Intelligence in Industrie 4.0', the impact of AI and the autonomy generated by it in the processes and results are mirrored in the application scenarios. An example of the effects of AI and autonomy is shown below in the scenario Value Based Services (VBS).

MANAGEMENT AND ACTION STRATEGIEST

'New knowledge and action options'

AI
- Collecting, evaluating, correlating of even unstructured information
- Creating hypotheses and finding of new correlations
- Symbolic deduction as important basis for automated decision templates

Domain experts
Technical experts
Users

Administration Shell
- Context and agreed semantics for all information and functions
- Explicit object dependencies re all asset structures
- Structured according to life cycles and application cases
- Structured information streams
- Objectives for application cases

'Basis of explicit knowledge'

Interaction Models
- Market place, negotiation and optimisation processes
- Implementation of complex process steps and processes
- Acquiring of new objectives

'Qualified action'

REAL WORLD

18   Plattform Industrie 4.0: Updated Application Scenarios of Plattform Industrie 4.0.
https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/fortschreibung-anwendungsszenarien.html.

The application scenario VBS shows how the value-added network can be designed in the service industry when the requisite product information and/or process information is available.

The idea behind the Smart Services World is to enable business enterprises (e.g. in the manufacturing industry) to make their data visible to select service providers. These in turn provide services that serve to refine the data to more universal forms of knowledge. On the basis of machine data, for example, maintenance instructions could be drafted or set-up plans optimised. The different services can access each other within the framework of a value-added network; what is created is an ecological system similar to a value chain.

There is no doubt at all that such services are going to profit in a major way from the ability of AI mechanisms to extract knowledge. The basic idea underlying many important AI algorithms is to transform input data into output data on the basis of patterns learned from the data available. For services that assist decision-making (servicing intervals, production plans, portfolio changes) or that detect anomalies (production flaws, technical defects, user conduct), these AI algorithms are extremely useful today already. With every analysis and its corresponding feedback from the service-user, the services can be improved and additional benefits generated for the user. Conversely, customers can use AI to create their own services or to make the best choices. Services are also able to learn from each other, which allows for the attainment of new levels of performance. The basis for all of this is qualitative high-grade data. One of the challenges here is the ability to separate

useful data from useless data, the latter comprising such things as measurement errors, non-representative random samples, or even sabotaged data. But in the long run, the services will improve. The result may even be unexpected synergies, where algorithms will cooperate with each other automatically without any human involvement. In such a scenario, it will no longer be a matter of only optimising the problems specified by human beings, but will involve a critical questioning of the assumptions underlying the model itself. Completely new solutions will be generated and offered as alternative options.

Security of Networked Systems

For the Hannover Messe, Working Group 3 – Security of Networked Systems – will prepare a publication with the title 'Artificial Intelligence (AI) in Security Aspects of Industrie 4.0'.

Legal Framework

Since 2016, Working Group 4 – Legal Framework – has been systematically identifying and working on the legal topics it deems relevant to Industrie 4.0 processes. The publication 'AI and Law in the Context of Industrie 4.0' will be published in time for the Hannover Messe 2019. The paper will deal with personality rights, access to and protection of data, liability issues, intellectual property, and employment law issues. It will also make recommendations for courses of action to be taken.

Illustration 4: Value Based Service (VBS).

Source: Plattform Industrie 4.0
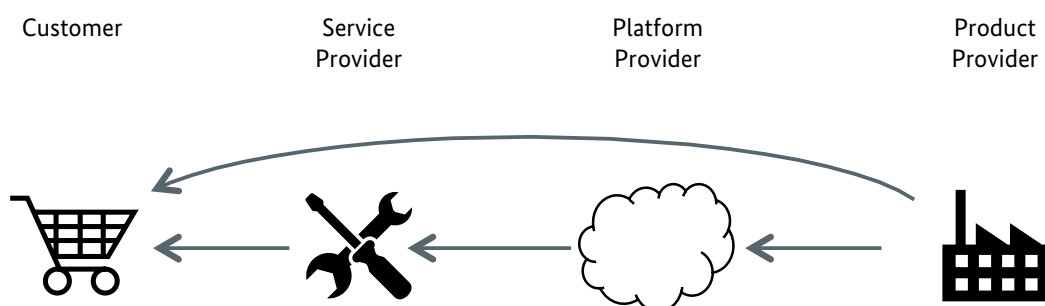
Work, Education and Training

Working Group 5 – Work, Education and Training – deals with the effects of Industrie 4.0 on human beings. Its primary focus is the structuring of the transformation of business enterprises in a way acceptable to the employer-trade union partnership, particularly through 'best practice' examples for the creation of framework conditions.

Working Group 5 is of the opinion that the use of AI will have a major impact on employment systems, regulations on working hours, workplace design, data protection, and on data security. AI will have a particular impact on the professional qualification requirements of workers, which the businesses must respond to with continuing education measures.

The group also wishes to examine the possible implications of the new business models that are made possible or inspired by AI (e.g. crowd employment).

AI represents a great deal of opportunity for German industry. It is important to act quickly. The timely provision of information and inclusion of all industrial stakeholders will guarantee a smooth and successful introduction of AI in the field.

Digital Business Models in Industrie 4.0

Working Group 6 – Digital Business Models in Industrie 4.0 – deals with the basic operating principles of digital business models in Industrie 4.0. The benefits of AI has become very evident in a variety of areas in the last few years. Most of these are found in the consumer world, one reason being the availability of large amounts and diverse forms of data in the B2C area. The use of AI technologies to exploit these large quantities of end-customer data has made the creation of completely new business models possible (e.g. solution for the 'long tail'). These new business models are often seen as the embodiment of what is at the heart of the digitised economy, namely a focusing on the customer's specific wishes and the resultant customising of products, coupled with no declines in the efficiency and quality of production. But even though such models and service platforms are well-entrenched and accepted on the market today in the consumer sector (B2C), such concepts will only become possible in industrial production (B2B) through the complete networking of production in Industrie 4.0 and through the rigorous use of AI methods. The automated compiling and processing of real-time data from industrial processes leads to new ways of optimising processes and of improving resource planning, and to a process-based use of personnel. Through an intelligent merchandise management, the classic ERP system (enterprise resource planning) is transformed into an intelligent and flexible planning tool. One goal could be the error-free production with the greatest possible compliance with production deadlines. The connecting of all production-proximate systems, such as status monitoring, production planning, or process quality, holds a great deal of promise. Such production-proximate data could be linked with commercial processes via interfaces. Optimised machine settings lead to improved energy consumption, to increases in capacity utilisation, to improved quality, and to the minimisation of raw-material usage. All of the combined data and information from process-related processes is compiled and analysed using an intelligent product planning and control system. Real-time processing allows for the detection of anomalies and problems. Optimisation possibilities can be recognised and directly responded to during the process work flows, all facilitated through the use of AI technologies. It is also possible to imagine hybrid AI systems in production in the future. These would analyse, in a cloud, the data from the various production facilities and would train the corresponding models. Local, production-proximate AI systems could then be periodically updated with the model parameters of the AI from the cloud in order to benefit from not only local forms of production optimisations but also from those effects that are only found in large amounts of data. To sum up, one can say that the use of AI systems will enable us to implement new business models more closely tied to customer needs, which will cause far-reaching changes in corporate structures.

Need for research on the industrial use of AI

Next to the legal issues and the technical-security issues, one of the biggest inhibitors to using AI technologies in industrial processes is the scarcity of data in relation to industrial applications. But this scarcity of data cannot be remedied by simply generating and storing larger and larger amounts of data. It has to be tackled through the use of algorithmic procedures. This is because the events that are relevant to industrial applications – defects in machinery, malfunctions in production processes, problems in relation to quality, or hazardous situations – actually occur quite infrequently. Empirical research shows that classic deep-learning procedures can only manifest their full powers upwards from a minimum number of example data sets. Research is therefore needed with respect to the inclusion of domain knowledge in machine learning processes, transfer learning, and methods of data augmentation for those data types particularly relevant for industrial use, i.e. especially for time sequences or signal data. It is also important that advancements continue to be made in the area of digitisation itself, and that it continues to become more effective, for example through the use of modern sensorics, better interpretability via self-describing data, modelling, cross-domain semantics, and flexible IoT communication infrastructures.

Another challenge is the accuracy and the performance capacities of the procedures needed for industrial applications. In the consumer sector, any recommendation system with an accuracy level of over 10% is already considered to be good enough. But even the simplest of industrial applications, such as optical quality controls or predictive servicing, require levels of predictive or detection accuracy of well over 90%. As to how these requirements are to be met, especially in light of the scarcity of data examples and the extremely unbalanced training data sets, is an unsolved issue in need of research.

The implementation of machine learning and artificial intelligence applications necessitates general expert knowledge in the area of data science (and long development periods) in addition to specific knowledge of the application domain. This renders many otherwise interesting industrial problems completely uninteresting for monetary reasons. A higher rate of automation and improved access to machine learning procedures are key prerequisites for a broad deployment across the various industries. AutoML is certainly a very interesting starting point for compensating for the lack of data scientists.

If artificial intelligence is to be used in security-critical systems, then reliability must be afforded a great deal of importance. The AI methods successfully in use today are often black box procedures whose behaviour and decisions can only be predicted or analysed with a great deal of difficulty. The intelligibility and comprehensibility of AI methods needs to be improved on the one hand, and security concepts for the use of AI in security-critical systems have to be developed on the other hand. AI is not primarily the solution, but is a procedure of systems engineering. Procedures from the area of software technology, such as model checking, are also potentially able to secure AI systems.

For the advancement of artificial intelligence in the industrial setting, there is a need for research in a number of areas. This includes such things as the intelligibility and comprehensibility of the methods used, ways of dealing with the absence of or too little data, and the development of methods with very low error quotas and methods with low configuration and engineering needs. All in all, research must serve to demystify AI, and it must do this by coming up with ways of making the applications of it intuitive, the decisions it makes comprehensible, and through the use of small amounts of data.

Other relevant areas of research include strategies for developing trust in industrial AI solutions, the answering of ethical questions regarding the use of AI in industry, and the best possible use of the connection between data-driven and physical models in AI.

AI in the civic and industrial debate

A variety of points need to be taken into account when discussing the ethical aspects[19] of the use of AI in industrial production in general and in the Industrie 4.0 in particular. It is first important to determine whether an AI system has a significant impact on human lives and/or on purely technical processes only. The discussion here definitely does not involve any notion of some kind of 'human-like intelligence' in the sense of a 'complete replacement of human beings by machines that have their own value systems and that autonomously define their own spheres of function.' The experts refer to this as 'strong AI', which has nothing to do with the AI mechanisms under review here. The kind of AI employable in industrial processes is a human-oriented AI that is aimed at supporting human capacities through automation processes.

The public discussion of AI in this form should be encouraged and conducted on a competent factual basis. This necessitates a high level of transparency with respect to both the chances and the risks of this technology and an estimation of the possible implications of it.

AI in industry

In the industrial setting, AI is used as a technological procedure aimed at improving processes and process flexibility, optimising costs, resources, time, energy, and machine performance, or at the automation of knowledge-based processes.

Because industrial AI applications are largely technical in nature and do not require intensive action on the part of human beings, it does not seem beneficial to talk about any kind of ethical purpose applicable to such technical applications. This does not mean, however, that these applications should not be subject to guidelines and regulation. There are already a wide variety of known industrial regulations in force, for example occupational safety regulations

---

19   According to the German dictionary 'Duden', ethics (in the vernacular usage) is defined as the 'sum of all moral rules and principles on which a [conscientious] attitude is based.' https://www.duden.de/rechtschreibung/Ethik#Bedeutung1a.

or functional safety regulations. These are founded on fundamental ethical principles developed over many years, which are basically European in their character and in use throughout the world.

A point worth mentioning with respect to the 'principle of autonomy' is that the decisions made by machines are typically made in adherence to task- and goal-oriented algorithms programmed by human beings. But on the other side, it is true that some parts of AI could consist of deep neural networks that are able to act in an unsupervised manner albeit within the bounds of narrowly defined parameters. The defined parameters are, however, specified by human beings, who therefore retain control over the results of the self-learning networks. Nonetheless, modern AI researchers regard 'explainable AI' as one of the areas that in the future will be capable of fostering trust in an autonomously acting AI.

Uncertainties in relation to the implications of AI applications

With respect to the concepts of traceability and auditability, it is important to consider the different contexts of the AI application. In the case of systems that impact human life in a significant way, lay persons should be able to clearly understand the causality of the algorithmic decisions made. But in the case of technical processes that have no impact on human life, then it is the experts who have to guide and

control the applications. Because of the fundamental changes the industrial sector is undergoing through the ever-increasing use of AI technologies, AI must be integrated as a discipline in educational training programmes and in life-long learning programmes in order to ensure that workers get the educational support they need.

In this respect, ethics must not become a negotiable factor in international competition; internationally recognised ethical guidelines applicable to the particular areas must be observed. These guidelines may apply to subject areas or to regional areas and may very well vary in their contents. Ethical conduct is ultimately based on trust and acceptance. Defining ethical guidelines for a trustworthy AI is a chance for Europe to offer the world a framework for all human-oriented AI applications.

The creation of guidelines is essential for the trust needed by businesses and in businesses when it comes to the acceptance of the digital transformation. Guidelines also serve as a means of preventing harmful AI scenarios within and outside industrial usage, which avoids the need for excessive bureaucratic and technical hurdles and keeps the costs associated with such hurdles (e.g. for assessment and monitoring systems in SMEs) to a minimum. The intention of Plattform Industrie 4.0 is to highlight the importance of human-oriented AI and to strike the right balance between a prudent approach to AI technologies on the one hand and the possible placing of too many constraints on its advancement on the other.

Summary of the AI debate to date

The first general recommendations for actions are in relation to the topics of acceptance, abilities/capacities, and international commercial usage. With respect to acceptance, a broad public debate on AI and the possible impacts of it is something to be encouraged. It should be conducted on a factually competent and transparent basis. Such a public debate allows for a discussion of the use and opportunities presented by AI and encourages people to think critically about the topic. It also serves the further modelling of AI methods and applications. Educational and professional training programmes must be set up in order to meet the current and steadily growing need for technological and methodical experts, and for application experts who are capable of transferring the methods to real applications. This should include not only experts and scientists but also the users and decision-makers as well. They must familiarise themselves with the behaviour of autonomous, AI-controlled processes, work with them, and perhaps even follow the instructions given by the autonomous processes.

Novel processes and alternative structural possibilities automatically give rise to a lot of questions involving the standards and regulations already in place. These standards and regulations often regulate planned and sometimes even certified procedures and systems (e.g. in the areas of functional safety and occupational safety) but do not yet provide answers to questions regarding the use of such

things as dynamic decision-making processes in AI systems. Also needed is backing for the further development of technologies and methodical knowledge. The focus of this should include the technological aspects of AI as well as a critical examination of sector knowledge, the sources of qualified data, and the industrial setting, such as AI usage in SMEs.

The use of AI is based on the ability to access information and qualified data. It involves a training phase on the one side, where the use of qualified data is of primary importance, and the operation of such systems using real-time and/or historical data on the other side. What is needed are methods for using training data and user data for commercial purposes (testing new business models and functionalities) and research projects (to further development AI). This requires a federal data landscape in which commercial enterprises and their customers come together applicatively, where business networks and consortiums can agree on common information and rules, and where both commerce and science can help themselves to information from a data pool of qualified data made available, for example, on a basis of trust. The political challenge here is the international networking and regulation of data use. What is needed here is the creation of internationally recognised guidelines for the use of AI and the data needed by it. In this regard, ethics must not become a negotiable factor in international competition.

AUTHORS:

Klaus Ahlborn, Gerd Bachmann, Fabian Biegel, Jörg Bienert, Prof. Dr. Svenja Falk, Prof. Dr.-Ing. Alexander Fay, Dr. Thomas Gamer, Kai Garrels, Dr.-Ing. Jürgen Grotepass, Dr. Andreas Heindl, Jörg Heizmann, Claus Hilger, Dr. Martin Hoffmann, Dr. Michael Hoffmeister, Michael Jochem, Johannes Kalhoff, Martin Kamp, Prof. Dr. Stefan Kramer, Dr. Bernd Kosch, Christoph Legat, Dr. Jan Stefan Michels, Dr. Alexander Mildner, Dr. Andreas Nettsträter, Rohitashwa Pant, Dr. Reinhard Pittschellis, Thomas Schauf, Dr. Hans-Jürgen Schlinkert, Dr. Marco Ulrich, Guido Zinke.