

International Center for Networked, Adaptive Production,  
Aachen

ICNAP Report 2024

# Foreword

---



**Change is the law of life and those who look only to the past or present are certain to miss the future."**

John F. Kennedy

Dear Digitalization Enthusiasts,

The year 2024 is marked by transition and continuity. Reflecting on community activities, we have experienced a successful year filled with insightful research and numerous gatherings. We take particular pride in having introduced new members to our fold, enhancing our collective journey towards interconnected and flexible production.

The Fraunhofer Institutes ILT, IPT, and IME in Aachen form the ICNAP Network, combining expertise in digitalization and Industry 4.0 with advanced infrastructure. The network features the Fraunhofer Edge Cloud, the Smart Manufacturing Network, over 2000 systems and machines for R&D in various technologies, and more than 28,500 m<sup>2</sup> of lab and hall space across the three institutes. Additionally, they access the 5G Industry Campus Europe, which includes indoor 5G networks and an extensive 5G network on the RWTH Aachen Campus.

This report offers a succinct recap of findings from five research projects that were voted on and conducted. The study titled "Zero Trust Architectures for Interconnected Industry" delivers in-depth knowledge about the protection of interconnected industrial systems. Moving on with our commitment to practical applications, "The Digital Twin Demonstrator – Bringing the Concept to Life" provides a tangible hardware and software demo that our associates can employ to gain practical insights for their individual use cases. The research "Seamless AI Integration through Plug&Produce Approach" strives to simplify the process for embedding and operationalizing AI in production machinery. In the same vein, "AI Everywhere - Generative AI for Production and Business Operations" investigates the wide-ranging impact of generative AI in both production and corporate settings. Lastly, "Towards a Dark Factory - Leveraging Multidimensional Twins in a Manufacturing Metaverse" discusses the use of comprehensive digital twins within a fully automated and self-sufficient manufacturing context.

Unlike prior reports, this document includes not just the outcomes of recent studies but also outlines ongoing initiatives in the ICNAP continuous working groups. The year 2024 focused on "Data-driven sustainability" and "Intelligent Sensing and Communication" as key themes. Through collaborative sessions in our community, we have crafted roadmaps for addressing these areas within ICNAP's framework.

I am honored to share that I have been appointed as Senior Executive Vice President for Research and Transfer, transitioning to the Executive Board of Fraunhofer in 2025. I eagerly anticipate embracing the new responsibilities and am committed to leveraging my international experience to propel innovation and further critical research and technology transfer initiatives at Fraunhofer.

Consequently, 2024 will be my last year serving on the ICNAP steering board, a position that has been both gratifying and pleasurable. As I move into my new role, I am confident in the team's ongoing excellence. I want to express my deep appreciation to the dedicated teams deeply passionate about the research and development at ICNAP, those who work with our clients and customers, and the team that manages and organizes ICNAP's daily operations, including the compilation and editing of this report. I am confident that their excellent work will be upheld in the future.

Yours sincerely



Prof. Constantin Häfner  
On behalf of the ICNAP Steering Board



# Table of Content

---

Introduction to ICNAP ..... 6

Zero trust architectures for interconnected industry ..... 16

The Digital Twin Demonstrator – Bringing the concept to life ..... 28

Seamless AI integration through Plug & Produce approach ..... 36

AI everywhere – Generative AI for production and business operations ..... 48

Towards a Dark Factory – Leveraging multidimensional twins in a manufacturing metaverse ..... 56

References ..... 64



# Introduction to ICNAP

## Networked, adaptive production

“What does Industry 4.0 mean for my production systems and what do I have to do to make my company fit for the future?” This is a question we frequently hear from manufacturers in Germany or anywhere else in the world. There is no simple, one-size-fits-all answer. It seems safe to expect, however, that increasing levels of digitalization in manufacturing environments and big data analytics will shape today’s manufacturing processes, subjecting these processes to radical changes which will make them leaner, sharper and more efficient.

only by joining forces between manufacturing industry, digitalization enablers and research, the potentials of Industry 4.0 can truly be exploited. Hence, ICNAP has been set up as a platform and community for collaborative R&D, to realize the visions behind “Networked, Adaptive Production”.

## ICNAP Topic Fields

### ICNAP Community key facts

ICNAP is a growing international community with partners from production technology, digital enablers and research institutions. Members work together to leverage the potential of digitalization. The Community offers network meetings, joint studies, discussion in working groups, corporate hackathon and more.

### ICNAP Research Partner

- Fraunhofer Institute for Production Technology IPT
- Fraunhofer Institute for Laser Technology ILT
- Fraunhofer Institute for Molecular Biology and Applied Ecology IME
- Manufacturing Technology Institute - MTI of RWTH Aachen University

Based on this request, the three Aachen-based Fraunhofer Institutes and other experts from industry and research have established the “International Center for Networked, Adaptive Production” (ICNAP) to find out which new approaches in information technology can lead the way towards Industry 4.0 and which requirements must be met. We are convinced that

ICNAP works in seven research areas, that cover the implementation of “Networked, Adaptive Production” and are shown in Figure 1.

### Sensor systems and data acquisition

How is relevant data along the entire process chain collected?

This topic field focuses on the collection and utilization of data in industrial and technical processes. One focus is on precise and efficient methods for data acquisition to optimize and control production processes. This data can be used, for example, for quality control and predictive maintenance. Various types of sensors, such as temperature, pressure, humidity or motion sensors and their applications are examined. A particular emphasis lies on the integration of these sensors into existing machines and processes, including hardware and software interfacing, to ensure smooth data transmission. Also important is the synchronization of multiple sensors to obtain consistent and temporally aligned data, which is essential for the analysis of complex systems. This will lead to new solutions for more efficient, flexible and intelligent industrial processes and will enhance companies’ competitiveness and innovative capacity.

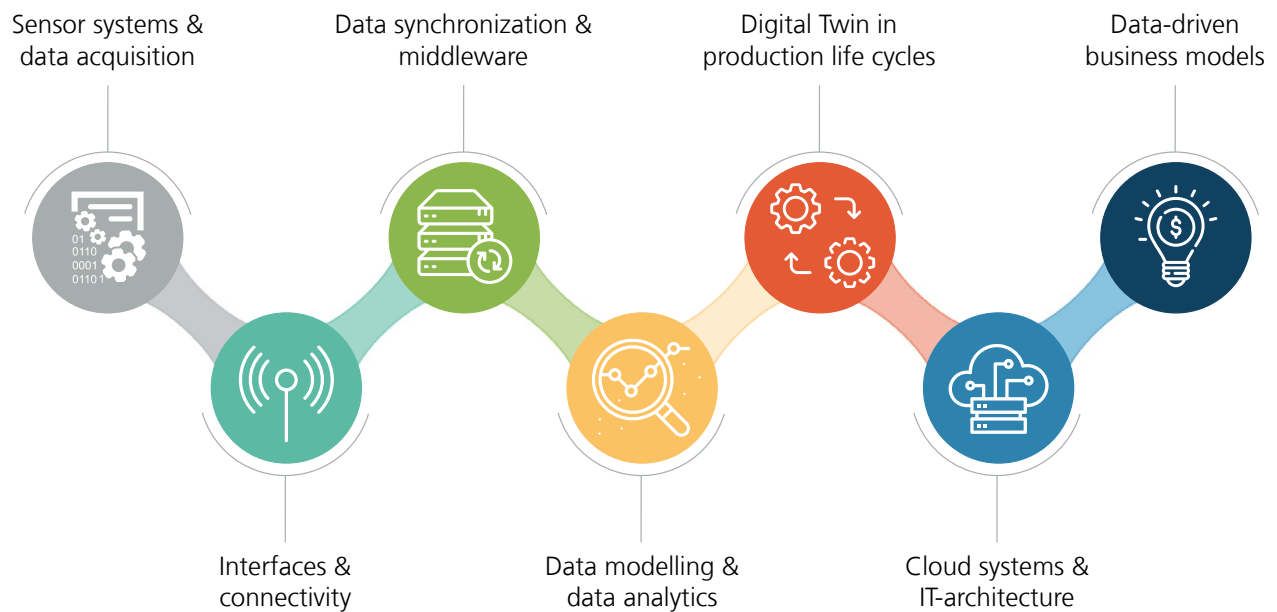


Figure 1: ICNAP topic fields.

## Interfaces and connectivity

How do the main components and systems communicate during the production process?

Once the production data is reliably recorded, a suitable communication protocol such as MQTT or OPC UA must be selected and a decision has to be made as to whether communication should take place via a wired or wireless connection. The ICNAP partners assess the data transfer requirements in terms of data throughput and latency for the production process in hand and can already issue recommendations at this point.

## Data synchronization and middleware

How are raw data (time-)synchronized and enriched with information?

The field of data synchronization and middleware mainly deals with the question how raw data is (time)synchronized between all entities of a communication network. This network usually includes sensors and computational hardware on the shop-floor, connecting desktop hardware for user/machine learning analysis and travelling upwards to MES and ERP systems. Information exchange between these various participants presents a major challenge due to the diversity of the single system. For example, on one side sensor systems act on analogue or digital voltage signals, which must be converted into digital representations, whereas ERP systems use document files or text-based communication. Additionally, each system implements its own

standards, communication protocols and security systems. Aligning these requirements often results in complex single-use-case solutions which are hardly adaptable or deployable on a general basis. The work within this topic field aims to find and/or develop uniform solutions based on the current level of technology. Therefore, one of the main tasks is to validate results of current research against the requirements presented by industries. Once the communication method has been selected and established within the production process, it is important to determine how to standardize procedures for filing data from various sources simultaneously and at various recording frequencies. The quality of different sources and the level of data accuracy also have to be defined at this point to accurately compare and evaluate the analysis results within the networked production environment.

## Data modeling and data analytics

How are the relevant data selected and what methods are used to obtain information from the data?

To provide employees in manufacturing companies – from production planners and quality managers to machine operators – with even better support for their decisions, it is important to define which process chain data is specifically relevant to them. This can be achieved by modelling the actual process digitally, based on the obtained data. Information can be generated from the data underpinning the structure of process knowledge using suitable methods of data analytics, especially machine learning algorithms or correlation analyses.

## Digital twin in the product life cycle

How is information and know-how digitally combined and visualized at different stages of the product life cycle?

Before the recorded information and the acquired know-how can be stored and used along the entire value chain, relevant data must be linked and related to each other. The result is a digital image of the real process which can be used to visualize the data in a goal-oriented and user-friendly way. Recommendations for action and feedback strategy for production can be derived from the knowledge obtained about the process.

## Cloud systems and IT architecture

How is the most suitable infrastructure selected and how are existing systems connected?

An efficient IT architecture is vital to the successful integration of digital tools in production. To this end, ICNAP tests data-bases and cloud systems for in-company or cross-company networks, taking into account current safety standards, norms and regulations.

## Data-driven business models

How are data-driven business models developed and integrated within the traditional product and service portfolio?

Not only does networked, adaptive production improve production by making additional knowledge available, it also paves the way for completely new forms of economic value creation through extended or new physical processes and products. Digital and data-based business models can be developed, classified and evaluated even for digital services such as machining-as-a-service or power-by-the-hour.

# Trend Topics for 2024

Networked, adaptive production can only be achieved by close collaboration between all the topics presented here. In 2024, the focus of the active working groups has been on two trending topics that are driving the community. On the one hand, aspects of intelligent sensing and communication are being elaborated, which enable safe and efficient data collection and transfer. On the other hand, motivation, challenges and enablers for data-driven sustainability are highlighted. The results from interactive sessions of the Summer Meeting in June 2024 will be presented here. These findings are the results of the dialog within the ICNAP Community and form the basis for future ICNAP activities.

## Intelligent sensing and communication

As part of the meeting, we conducted a workshop on the trend topic "Intelligent sensing and communication". In this workshop, we addressed key aspects of modern sensor technologies and communication systems. Our focus was on adaptive systems that can dynamically adjust to changing conditions, intelligent technologies that utilize data analysis

and machine learning to be self-learning and decision-capable, and networked solutions that enable seamless communication and interaction between various devices and systems. We also talked about the importance of resilient systems that can resist disruptions and failures, as well as security mechanisms that ensure the protection of data and systems against unauthorized access and cyber attacks. Finally, we looked at sustainable approaches that aim for efficient resource utilization and long-term environmental friendliness.

In a first workshop session, we collected input on each of these topics. From these inputs, we derived four focus areas: "Security by Design", "Frugal Communication & Sensing", "Resilient Infrastructure" and "Standardized Interoperable Communication". (Figure 2) We conducted discussion rounds to identify which initiatives are particularly important for implementing the issues discussed. Subsequently, we created a roadmap for these four topics. The aim of our workshop was to gain an initial overview of these topics and to identify areas of high importance for the ICNAP community. The results of our discussions and the roadmap for further action are described in the following.



## Security by design

The term "Security by design" refers to the approach of integrating security into the development process of systems and applications from the outset, rather than adding it afterwards. The aim is to minimize vulnerabilities and enhance resilience against potential threats. Several important research topics were identified during the discussion. First, the automation of data classification, which involves developing automated methods to determine the protection needs of data. Second, secure communication protocols and methods between machines and various communication platforms, to ensure the security of machine communication. Third, the implementation of security measures directly on the shop floor to protect industrial processes.

Another key topic was the automation of risk analyses to identify security risks more quickly and accurately. Additionally, two databases were proposed: a comprehensive database of known security risks and a database of appropriate mitigation measures. Finally, the requirements and best practices related to the NIS2 Directive and the Cyber Resilience Act (CRA) were discussed to meet regulatory requirements and strengthen cyber resilience.

## Frugal communication and sensing

For "Frugal communication and sensing", the motivation is clear: making communication and sensing more sustainable. This is based on the mindset of only capturing and transmitting data when it's useful and creates added value.

Applying this concept is possible on various levels as discussed in the meeting: At sensor level configurable sensors allow for flexible installations to adapt to various environments and requirements. The sense-on-demand approach focuses on collecting data only when it is necessary and useful. Based on conditions and triggers, the sending interval can be adjusted. Metadata including timestamps and data classification as well as encryption also become relevant as they highly influence the communication resource and energy needs. A selection process of relevant data, sending intervals, necessary metadata and encryption, is needed to support minimizing energy consumption and reducing the carbon footprint.

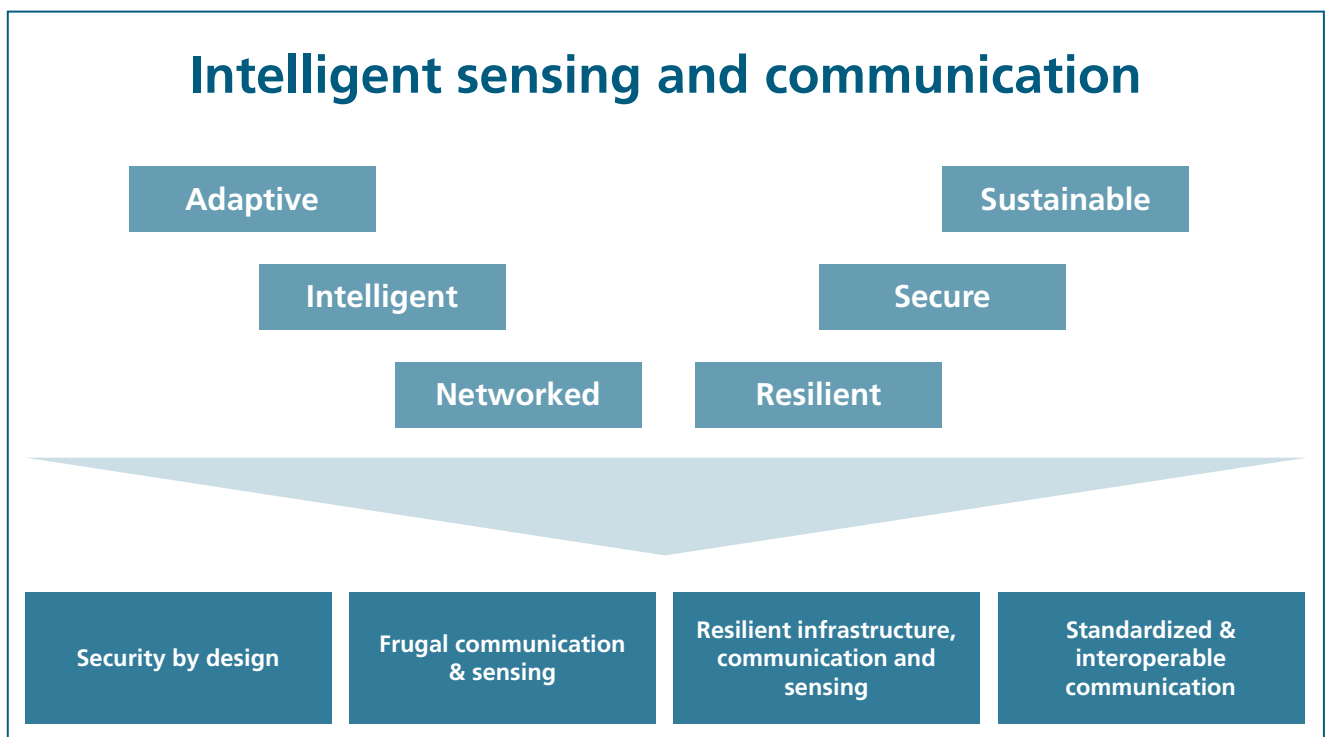


Figure 2: Intelligent sensing and communication.

## Resilient infrastructure, communication and sensing

Resilient infrastructures for communication and sensing are crucial for maintaining reliable and secure operations in modern systems. An important consideration is the choice between wired and wireless communication. Wired communication offers higher stability and security, making it ideal for environments where interference or unauthorized access is a concern. In contrast, wireless communication provides flexibility and ease of deployment, though it may be more vulnerable to disruptions. Anomaly detection plays a vital role in identifying and mitigating potential threats to the system's integrity. By continuously monitoring data patterns, the system can detect irregularities that may indicate faults or malicious activities. This proactive approach is vital for maintaining the resilience of the entire system.

Evaluating the resilience of a communication and sensing system involves assessing its ability to withstand and recover from disruptions. This includes testing the system under various stress conditions and ensuring that both communication and sensing components can operate effectively even when some parts fail.

Data integrity is maintained through robust encryption, redundancy, and error-checking mechanisms. These measures ensure that data remains accurate and unaltered throughout its transmission and storage, even in the face of potential system failures or attacks.

Finally, the ability to swap sensors without disrupting the system is crucial for long-term resilience. This requires modular designs and standardized interfaces, allowing for quick replacement or upgrading of sensors, ensuring continuous operation and minimizing downtime.

By discussing these topics during the meeting, three general aspects for further projects have emerged. First developing a method to evaluate the resilience of the whole system, that includes sensor hardware, communication channels, processing boards and the monitoring systems. Second would be the use of the Kalman filter technology to detect interruptions via state estimation and/or anomaly detection for sensor faults. The last idea would be to present a hot-swap approach for the sensor itself, which reduces downtime due to repairs.

## Standardized and interoperable communication

Standardized and interoperable communication is essential in modern industrial and technological environments, facilitating seamless data exchange between various systems and components. This approach needs to increase the focus on the standardization of information models and databases, ensuring that data is structured and interpreted uniformly across different platforms. These standards enable diverse systems to understand and utilize shared data effectively.

In the middle layer, standardized protocols for data collection and communication are crucial. These protocols act as a bridge, allowing different devices, applications, and systems to communicate without compatibility issues. This layer often includes the use of standard Manufacturing Execution System (MES) interfaces, which ensure that production data from various sources can be integrated and managed effectively. Similarly, standardized sensor integration allows for consistent data input from various sensors, regardless of their manufacturer.

When it comes to cloud architecture versus closed on-premise networks, each has its pros and cons: Cloud architecture offers scalability, flexibility, and ease of access from any location, making it ideal for companies looking to manage and analyze large datasets. However, it can pose security concerns and requires reliable internet connectivity. On the other hand, closed networks, which are entirely on-premise, offer enhanced security and control, with data staying within the local infrastructure. This setup is often preferred in industries with strict data protection regulations but can be less flexible and harder to scale compared to cloud solutions.

In summary, standardized communication protocols, alongside interoperable interfaces and models, are foundational for integrating various industrial systems, enabling both cloud-based and on-premise architectures to operate effectively depending on the specific needs and constraints of an organization.

# Data-driven sustainability

During the summer meeting a community workshop on our second trend topic “Data-driven sustainability” was conducted. Sustainability, especially in the context of manufacturing and production, is gaining more and more traction. In the automotive sector, numerous pilot projects and first implementations of the soon required battery passport can be seen [1]. As ICNAP we decided to approach the topic mainly in the following three spotlight areas:

**Circular economy:** Circular economy (Figure 4) in manufacturing refers to a system, that minimizes waste and maximizes resource efficiency by designing products for reuse, recycling, and repair, creating closed-loop production cycles [2]. We can already find a lot of circular approaches, for example when it comes to the recycling of metal car components. But often a significant loss of previous value creation is implied, which results in down-cycling, instead of the aspired recycling (Figure 3).

**Life-cycle assessment (LCA):** Life-cycle assessment (Figure 4) refers to the evaluation of the environmental impacts of a product or process through its entire life cycle, from raw material extraction to disposal or recycling [4]. One of the most notable projects in regard to LCA currently being pursued by major automotive OEMs is the battery passport. This is a digital

record, that tracks the environmental and material footprint of an electric vehicle battery through its lifecycle, ensuring transparency and sustainability from production to recycling.

**Supply chain sustainability:** Supply chain sustainability (Figure 4) in manufacturing refers to integrating environmental and socially responsible practices across the entire production process, from sourcing raw materials, to delivering finished goods. It involves minimizing environmental impact, reducing waste, and ensuring ethical labor practices. Current trends in supply chain sustainability include a focus on reducing carbon emissions, adopting circular economy practices (recycling and reuse of materials), as well as increasing transparency through the usage of digital tools like e.g. the digital material twin for the optimization of sheet metal processes, that has been presented at our 2024 Summer Meeting.

Resulting from the discussions with the ICNAP Community members, two main sub-topics were identified, that play an important role for successful implementation in the three spotlight areas: “Reporting and regulation” and “Supplychain data sharing”.

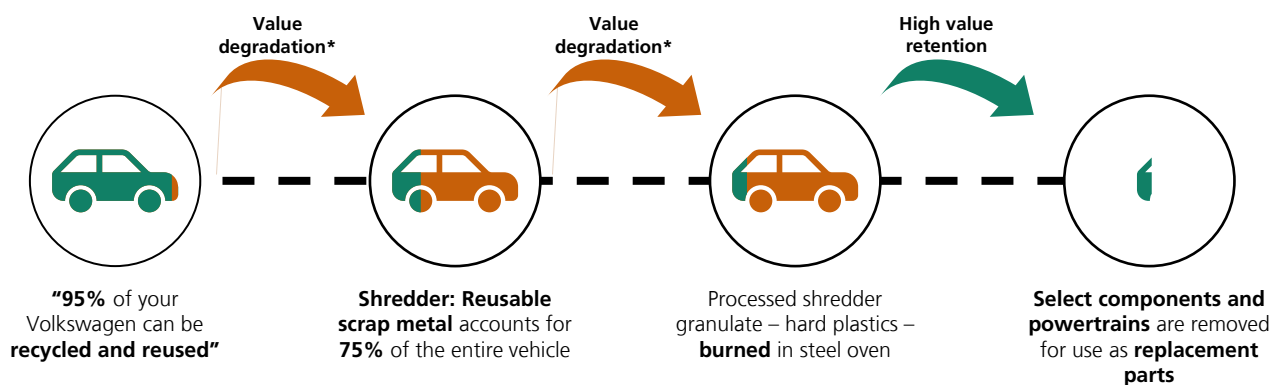


Figure 3: Value degradation in a typical car recycling process [3].

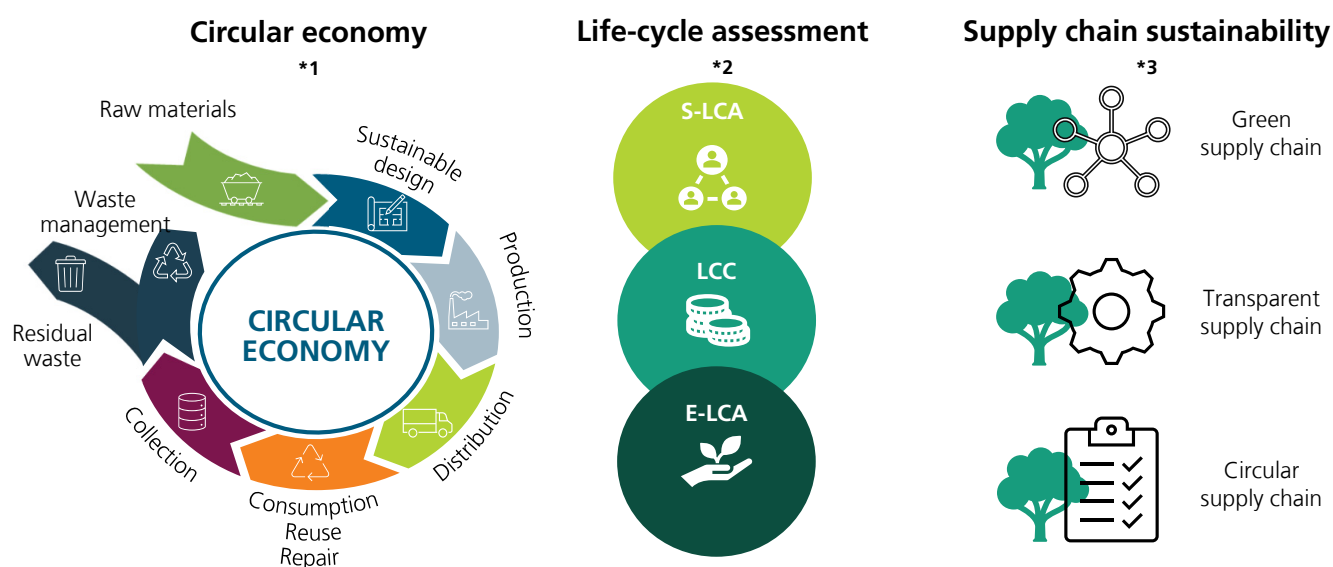


Figure 4: Pillars for data-driven sustainability.

## Reporting and regulation

Reporting and regulation are essential in all three major focus areas discussed. During the on-site workshop, the community highlighted several challenges they face concerning these topics.

Interestingly, the main difficulties are not related to any particular law or regulation, but rather to the complexity of navigating the interconnected and often fragmented regulatory landscape. Regulation and certifications vary widely across regions, and even within the same region, requirements can overlap. For example, complying with product-specific regulations like the EU directives on batteries or single-use plastics often intersects with broader directives such as the Corporate Sustainability Reporting Directive (CSRD). Identifying these overlaps and ensuring compliance with the various requirements is a major challenge that requires a well-coordinated and efficient approach.

In addition to this overarching challenge of complexity, the community also faces operational hurdles. For reporting purposes, companies need to collect and compile data from multiple sources, ranging from internal departments to different

external suppliers. This can lead to high manual effort in data evaluation and documentation – and falls short of building a high-quality data foundation. A key issue here is the lack of standardized frameworks. This gap extends both to technical aspects, such as standardized information models, and to broader factors, such as unified contracts or agreements for data exchange. Addressing this lack of consistency is crucial to improving reporting efficiency and regulatory compliance.

Regulatory requirements outside the sustainability context further contribute to the complexity mentioned earlier. Recently enacted regulations, such as the EU AI Act and the new Cybersecurity Act, introduce additional challenges, particularly concerning data generation and utilization. Moreover, there remains a considerable degree of uncertainty regarding how these regulations will be implemented and enforced, as there is currently limited case law or precedent to guide compliance efforts.

In conclusion, although reporting and regulation may seem tedious at times, they are set to play an increasingly significant role for the ICNAP community. Some challenges are already evident, while others are gradually emerging on the horizon (Figure 5)

## Supply-chain data sharing

Sharing data along the supply chain is an enabler for many sustainability topics and therefore directly linked to the three focus areas mentioned above. However, in practice, exchanging information often happens on an ad-hoc basis that lacks a coherent or scalable system. Frequently there is no data exchange at all, even though doing so could provide tenable benefits for the parties involved. The reasons for this are many-fold and differ case by case, but they can be broadly summarized into a few categories.

One category for obstructing causes springs from a human factor. Simply put, data is not being exchanged along the supply chain because people “don’t want to”. For example, the reason could be that there is a lack of incentive for a business to share its data, or management support for a systematic exchange is missing. In many places obligations to keep data secret (especially when it comes to production data and personal data) and competition concerns prevent collaboration between supply-chain participants.

A second category are technical barriers: Often it is unclear what form a systematic information exchange could take. Additionally, a lack of standardized communication and interoperability between participants increases complexity

and ultimately costs for data sharing. Moreover, given the often-sensitive nature of supply chain data, prospective data providers want to make sure that their information is only accessed according to agreed rules. How these rules can be enforced is then frequently unclear.

Lastly some issues also originate from a high barrier of entry. In order to satisfactorily share data, many preconditions have to be fulfilled. Not only does any exchange have to be secure and confidential, but it also has to be transparent (i.e. open about the data’s source, context, provenance, limits, etc.). The need for extensive data classification (e.g. whether the data is related to persons, machines, and so on) is furthermore a factor that raises costs and complexity.

In summary, several factors and causes (Figure 6) lead to data sharing along the supply-chain being not as common as would be useful for many sustainability goals. However, these problems are addressable and a growing need and motivation for green production is set to motivate the ICNAP community to tackle them.

### Fundamental challenges

Certifications for different countries	Standardization in reporting
Interoperability between different regulations/standards	

### Challenges data aggregation and other

Report compilation (multiple sources, e.g., suppliers)	Lack of standardized information models
Contracts/agreements for data exchange	Challenging to make “profitable”
AI-regulation hard to get an overview	Certification cyber security – who can do it? How?

Figure 5: Challenges of reporting and regulation.

### Problems:

No standardized communication (increases complexity)	Competition prevents sharing
No systematics of exchange	Lack of incentives (could increase chance of sale)
Unclear how data access rules can be enforced	Lack of management support
Export regulations impose more obstacles	Personal data is especially sensitive
Legal contractual obligations stand in the way	Lack of interoperability
Need to keep production data secret / internal	

### Preconditions:

Transparency	Data confidentiality
Trustability	Consensus (among senders / recipients)
Secure exchange	Confidentiality (need-to-know)
Data classification (e.g. personal, machine-related, etc.)	

Figure 6: Problems and preconditions for supply-chain data sharing.

# Conclusion

The trend topics of data-driven sustainability and intelligent sensing and communication lead to fruitful interaction with the community during the Summer Meeting emphasizing their importance for the future. Discussions uncovered the necessity of breaking down trend topics into more specific sub-topics. These insightful conversations may lead to study proposals and bilateral projects aimed at advancing these key areas.

Each year, the community selects and discusses new topics in collaborative studies. The studies are conducted in three phases: a pilot study, a detailed study and the application and business cases. The goal is to develop a hands-on demonstration of the applicable technologies and methods within the

ICNAP community, either on the ICNAP shop floor or at one or more of the member companies. The studies can be cross-topic covering one or more ICNAP topic fields. In the following chapters, we provide a summary of the key results from the five studies conducted in 2024.





# Zero trust architectures for interconnected industry

---

**Max Ortmann**

Research Fellow

Digital Infrastructures

Fraunhofer Institute for Production Technology IPT

**Prof. Dr.-Ing. Robert H. Schmitt**

Member of the board of directors of Fraunhofer IPT and holder of the chair for Production Metrology and Quality Management at the WZL | RWTH Aachen University

# Introduction

The manufacturing sector has witnessed a transformative shift in recent years, driven by the rapid adoption of digital technologies. This paradigm shift, often referred to as Industry 4.0, has resulted in the seamless integration of physical and digital systems, leading to enhanced operational efficiency and growth. While these technological advancements offer numerous benefits, they also introduce significant cybersecurity risks, which are constantly increasing due to a multitude of vulnerabilities in digital elements and growing global tensions. Data theft, industrial espionage, and sabotage caused a total loss of 205.9 billion euro in Germany alone in 2023, and this trend is on the rise [5]. In 2024, both analog and digital attacks are projected to increase by approximately 29 %, reaching a total of 266.6 billion euro [6]. The alarming trend observed in Germany is mirrored on a global scale, with the cost of cybercrime anticipated to escalate to 13.82 trillion Dollar (approximately 11.73 trillion euro) by 2028, up from 9.22 trillion Dollar (approximately 7.83 trillion euro) in 2024 [7].

## Emerging threats and challenges in Germany's security landscape

The recent study by Bitkom e.V., the industry association for the German information and telecommunications sector, surveyed 1,003 participants from various industry interest groups [6]. The findings reveal that 8 out of 10 companies have experienced cyber attacks, and two-thirds of these companies feel that their very existence is at risk. Cyberattacks have risen by 7 % compared to the previous year, particularly targeting operational processes, information, and production systems. According to the surveyed companies, 70 organizations from German economy were affected or suspected of being affected between June 2023 and June 2024. Simultaneously, physical attacks have also surged, with a 15 % increase in the theft of physical documents, personnel files, patents, machines, and components compared to the previous year. Additionally, there has been a 13 % rise in the interception of meetings and phone calls on-site over the same period.

## Enhanced security and regulatory requirements for smart factories

The global increase in cyberattacks, which are mainly attributed to state actors and organized crime, is causing the production sector to be increasingly targeted. As digitalization advances and the interconnection of operational technologies (OT) expands, not only IT systems but also various aspects of manufacturing companies could become the prime targets for cyberattacks. The integration of modern technologies such as 5G, AI, cloud systems, and smart sensors into traditionally isolated fieldbus systems is driving a convergence of information technology (IT) and operational technology (OT). This convergence significantly broadens the attack surface in production environments and creating new security vulnerabilities, especially when legacy systems are integrated into digitalized production environments. To summarize, according to a study by Allianz SE, companies see the greatest risks in data breaches, cyberattacks on critical infrastructure and physical assets, the increase in malware/ransomware attacks and disruptions due to the failure of digital supply chains, cloud and service platforms [8].

In response to these escalating threats, the European Union has implemented a series of regulations aimed at enhancing cybersecurity across various sectors, including manufacturing. Regulations such as the Network and Information Security (NIS2) Directive, the EU regulation on machinery and the Cyber Resilience Act (CRA) seek to establish a unified approach to cybersecurity, mandating that organizations adopt specific measures to protect sensitive data and ensure the resilience of critical infrastructures throughout the whole supply chain. Compliance with these regulations is not only a legal obligation but also a crucial step in building trust with customers and partners, reinforcing the need for robust security practices in production environments.

## **Embracing secure digitalization: Zero Trust as a key strategy for future-proof manufacturing?**

One promising approach to addressing these challenges is the adoption of a Zero Trust security model, particularly in digitalized Operational Technology (OT) settings. The Zero Trust framework operates on the principle of “never trust, always verify”, meaning that no device or user is trusted by default, regardless of their location within the network. By implementing strict identity verification protocols and continuously monitoring network traffic, manufacturing companies can significantly reduce the risk of unauthorized access and lateral movement within their systems. This approach is particularly effective in OT environments, where traditional perimeter defenses may no longer suffice due to the convergence of IT and OT systems. By fostering a culture of continuous security vigilance, the Zero Trust model can help mitigate the risks posed by cyber threats, ensuring the integrity and availability of critical production processes. However, implementing a Zero Trust architecture involves significant effort and high complexity. Additionally, the lack of harmonized standards for interoperability of product functionality presents a major challenge for companies. For this reason, the study “Zero Trust Architectures for Interconnected Industry” addresses the question:

“How can Zero Trust be established in OT environments considering industry specific requirements and is it reasonable?”

The first section outlines the obligations imposed on manufacturing companies by the European Union within the framework of cyber regulation. Based on these requirements, a comprehensive study on Zero Trust is conducted, with a particular focus on the adaptation of its concepts in digitalized operating environments. Finally, the study offers a best practice guide, providing companies with a step-by-step overview of the implementation of Zero Trust in production settings, while identifying associated opportunities and risks.

# Cyber regulation in the European Union

Due to the rising number of cyberattacks on companies within the European economy and critical sectors, the European Union is enacting an increasing number of cyber regulations, now also impacting the manufacturing industry for the first time. Notably, the Network and Information Security (NIS2) Directive, the EU Machinery Regulation, and the Cyber Resilience Act (CRA) establish a regulatory framework that targets the processes of manufacturing companies in the EU, as well as hardware and software products distributed within the European Single Market. This legal framework not only binds individual companies within the EU but also extends to suppliers, thereby encompassing the entire supply chain. Consequently, companies that manufacture products outside the EU and sell them within the EU are also subject to these regulations. A brief overview of the relevant regulation is given in Figure 7.

The following sections provide a detailed examination of the NIS2 Directive and the CRA and summarize the obligations for companies.

## Network and Information Security (NIS2) Directive

With the implementation of the Network and Information Security (NIS2) Directive, which must be transposed into national law by October 18, 2024, the European Union sends a clear message. This directive broadens the definition of critical security sectors to include manufacturing companies as important entities, such as those in the machinery sector, with a workforce of 50 or more employees or an annual revenue exceeding 10 million euro. It is estimated that around 30,000 businesses within Germany will be affected directly by these measures, with approximately 80 % of these companies being unaware of their obligations.




NIS2 Directive	Regulation on Machinery	Cyber Resilience Act
<div></div> <div><b>Focus on compliance &amp; governance</b></div> <div>Obligations in governance, awareness, risk management, reporting obligations and risk management in the supply chain.</div>	<div></div> <div><b>Focus on safety &amp; security of machinery</b></div> <div>Regulations for safety and security of machinery and industrial control systems (ICS), including risk management and implementation of security measures.</div>	<div></div> <div><b>Focus on product security for hardware &amp; software</b></div> <div>Key objectives: “security by design”, risk management, reporting of vulnerabilities and incidents and transparency.</div>
<b>Comes into force on October 17, 2024</b>	<b>Comes into force on January 20, 2027</b>	<b>Expected to come into force in November 2027</b>

Figure 7: Overview of upcoming cybersecurity regulations that affect the manufacturing sector.

Production companies that fall under the NIS2 Directive must comply with certain obligations. These obligations, outlined in the NIS2 Directive, include:

- **Governance:**

Management bodies are responsible for approving risk management strategies, monitoring their implementation, and being held accountable for any violations that occur.

- **Awareness:**

Regular security training should be conducted to enhance knowledge and skills at all levels, including management, to effectively identify risks and apply cybersecurity procedures.

- **Risk management:**

A comprehensive approach to risk management should include risk analysis, incident handling, business continuity planning, security measures for network and information systems, the use of cryptography, multi-factor authentication (MFA), and effective asset management.

- **Reporting obligations:**

Organizations must report any significant security incident within 24 hours of becoming aware of it. An assessment of the incident's severity and impact should be provided within 72 hours, with a final report due within one month.

- **Supply chain:**

It is essential to ensure the security of supply chains by addressing both technical and non-technical risk factors.

As a result, the company's management is directly accountable for the implementation and oversight of cybersecurity measures. Additionally, technical, operational, and organizational measures must be established to manage risks, aiming to prevent, detect, and respond to potential cyberattacks. The scope and scale of these measures should align with various evaluation factors, such as company size, incident likelihood, and the societal and economic impact of security breaches.

In the event of significant security incidents causing serious operational disruptions, affected companies must fulfil reporting obligations. An initial warning must be reported within 24 hours of becoming aware of a significant security incident, followed by a comprehensive assessment within 72 hours, including severity and impact ratings. Interim reports on relevant status updates may be requested, with a final report due no later than one month after the incident. These reporting obligations also apply if third parties or institutions may suffer significant material or immaterial damages. Notably, NIS2 addresses supply chain risk management, acknowledging that attackers may exploit trust within the supply chain to introduce malicious components or compromise its integrity.

For manufacturers affected by NIS2, non-compliance with these obligations can result in fines of up to 7 million euro or 1.4 % of the total global revenue from the previous financial year. Authorities are authorized to conduct on-site inspections and oversight measures, which can also be delegated to third parties, including trained external professionals. In cases of violations, company executives may be held personally accountable.

## Cyber Resilience Act (CRA)

The Cyber Resilience Act (CRA) addresses a critical gap in cybersecurity by focusing on the security of products entering the European market, an area that often receives less attention despite investments by companies in their own network security measures. To enhance both corporate governance and product security, the NIS2 Directive and the CRA complement each other effectively. However, while the NIS2 Directive is set to come into force end of 2024, companies will have additional time to fully implement the requirements that are mandatory under the CRA. The CRA was adopted by the EU Parliament in March 2024 and is awaiting ratification by the EU Council at the time of writing. It is anticipated that this ratification will occur before the end of 2024, initiating a 36-month transition period. As a result, effective implementation of the CRA is expected by November 2027.

Its primary objective is to ensure that digital products and services are designed with robust security measures from the outset, thereby reducing vulnerabilities that could be exploited by cybercriminals. One of the key components of the Act is the establishment of specific security requirements that manufacturers must adhere to when developing digital products. This includes ensuring that products are resilient to cyber threats and that they incorporate security features that can withstand potential attacks. Transparency is another critical aspect of the Cyber Resilience Act. Manufacturers are required to provide clear and accessible information regarding the security characteristics of their products. This includes details about potential risks, security updates, and how users can protect themselves. By fostering transparency, the Act aims to empower consumers and businesses to make informed decisions about the digital products they use. In addition to these requirements, the CRA mandates that companies report significant cybersecurity incidents to relevant authorities. This incident reporting obligation is designed to facilitate a swift response to cyber threats and to improve the overall understanding of the cybersecurity landscape within the EU. By collecting data on incidents, authorities can identify trends and develop more effective strategies to combat cybercrime.



The key provisions of the Cyber Resilience Act (CRA) can be summarized as follows:

- **Risk assessment and support throughout the entire product life cycle:**

Manufacturers must conduct risk assessments to identify potential vulnerabilities in their products and services. They are then required to implement appropriate security measures to mitigate these risks. In addition, manufacturers must provide free security updates for at least 5 years.

- **Security by design and by default:**

Products and services must be designed with security as a core consideration from the outset. This means that security features should be built into the products rather than added as an afterthought. Additionally, products must be configured to have secure default settings.

- **Reporting of security vulnerabilities:**

Manufacturers and importers are obligated to report cybersecurity incidents to the relevant national authorities.

- **Product labeling and transparency:**

Products must be labeled with information about their security features and any known vulnerabilities. Consumers will have the right to be informed about the security implications of the products they purchase.

Noncompliance with the CRA can impose substantial sanctions. Violations of fundamental requirements can result in fines of up to 15 million euro or 2.5 % of a company's total global annual revenue, whichever is higher. For breaches of other obligations, fines can reach 10 million euro or 2 % of global revenue. Additionally, providing false or misleading information to notified bodies can lead to fines of up to 5 million euro or 1 % of total global revenue.

# Zero Trust as a security strategy in manufacturing

The implementation of security measures involves significant complexity and effort. The production sector presents a unique challenge, as it must consider both IT and OT systems. As these areas converge through digitalization and networking, the potential attack surface for companies expands, making them more vulnerable to threats, even at the field level [9]. A minor oversight can create a critical vulnerability, allowing unauthorized access to internal systems. Therefore, adapting cybersecurity strategies is essential for organizations to effectively plan, assess risks, and monitor their network and information systems. With forthcoming cybersecurity regulations, compliance with established standards may soon become mandatory.

Numerous recognized national and international norms and standards are documented in the literature. In a globalized economy, harmonized standards are especially beneficial, as they are acknowledged across various countries and increasingly serve as prerequisites for trade between institutions. For the manufacturing industry, ISO/IEC 27001, which focuses on information security management systems, could help with compliance with the NIS2 Directive, as ISO 27001 defines measures for implementing the minimum requirements of the NIS2 Directive. For example, through the introduction of an information security management system (ISMS), processes can be introduced for risk management or for handling security incidents. The prior introduction of a quality management system in accordance with ISO 9001 is recommended. IEC 62443 can be used in a similar form for the implementation of the requirements by the CRA in the production sector. In particular, IEC 62443-4-1 and IEC 62443-4-2 define technical and organizational requirements for "Security by Design".

While many standards share common elements, they often have distinct certification requirements, especially in the manufacturing sector. This can pose challenges for organizations seeking compliance, as they must navigate the unique criteria and processes of each standard. Furthermore, differing interpretations of similar concepts can result in inconsistencies in implementation across manufacturing environments. However, obtaining security certification does not guarantee that a company is fully protected against cyber attacks. Often, only minimum requirements are mandated, leading to the implementation of isolated solutions that may not address broader security needs once certification is achieved. To address these challenges, the concept of Zero Trust is gaining prominence. Zero Trust is a strict security strategy that companies can adapt to improve technical and organizational protection and ensure compliance with cyber regulations. The importance and future viability of Zero Trust as a security concept is underlined by the requirement for all US

government agencies to convert their infrastructures to a Zero Trust architecture. The following section outlines the fundamental principles of Zero Trust and presents best practices for its application in manufacturing environments.

## Core principles of Zero Trust

Zero Trust is a strict security paradigm that assumes a breach has already occurred. It operates on the principle of least privilege, requiring all entities, for example devices, user or systems, to prove their identity and authorization before accessing resources [10], [11]. This eliminates implicit trust among entities and transforms traditional perimeter-based security into a multi-layered, integrated security approach. Communication between entities necessitates explicit verification and earned trust through reliable evidence. This continuous trust check minimizes the risks to confidentiality and integrity but can impair availability. It should be noted that the availability of resources in particular is essential for production. In summary, Zero Trust is defined by three core principles:

### 1. Assume breach:

There is no longer a distinction between internal and external networks; the internal network is always considered insecure, and trust is never granted permanently. Trust is continuously assessed based on dynamic access policies, ongoing monitoring, and risk analyses, with access decisions made anew each time.

### 2. First verify, then trust:

The absence of implicit trust necessitates that every entity must authenticate and be authorized to access resources, with strong authentication playing a crucial role.

### 3. Least privilege:

The principle of least privilege means that only entities requiring access are granted it. This requires resources to be divided into smaller units and permissions to be assigned as granularly as possible. A smaller access radius limits uncontrolled data exfiltration, data manipulation, and lateral movement in the event of malicious access.

## Zero Trust reference architecture

The principles of Zero Trust do not dictate a specific architecture and remain unstandardized [10], [11]. Instead, guiding frameworks are provided for processes, identities, system design, and their interactions. For example, the NIST Special Publication 800-207 "Zero Trust Architecture" provides a reference architecture that is also supported by the German Federal Office for Information Security (BSI) in a "Zero Trust" position paper from 2023 [10]. The reference model and the logical components of a Zero Trust architecture are outlined in the following and shown in Figure 8.

The access decision functionality in a Zero Trust architecture is referred to as "Policy Decision Points" (PDP). The PDP component ensures that access requests are valid. It can be a local entity within the organization or an externally hosted service. For evaluation, it could utilize the organization's access policy and could gather information from various sources to assess trustworthiness. If the trust assessment is validated, the PDP could issue a restricted access permission from a device, user or system to the "Policy Enforcement Point" (PEP). The PEP then enforces the decision made by the PDP. To ensure the integrity of the communication paths, there should also be a separation between the communication required to control and configure the internal network and the communication used for application access. According to the Federal Office for Information Security (German: Bundesamt für Sicherheit in der Informationstechnik, abbreviated as BSI), a physical separation

should take place for requirements with increased protection needs, whereby a logical separation is also sufficient for standard requirements [12]. In the Zero Trust reference model, this is termed the "Control Plane" and the "Data Plane". The process is illustrated in Figure 8.

The specific measures are largely influenced by the company's structure and may include factors such as IP address ranges, geographical access distribution, time-based access controls, the use of certificates, or hybrid dynamic models.

## Best practices for adopting Zero Trust in digital production environments

Implementing Zero Trust in digitalized and networked production environments necessitates a well-planned change management process with tailored measures. Understanding the current status of security implementations is crucial. The Zero Trust Maturity Mode (ZTMM) serves as a valuable framework to guide the effective integration of Zero Trust principles [11]. The following section therefore outlines a best practice guideline based on the reference architecture and the ZTMM that focuses on the production sector.

The ZTMM defines seven principles for the successful implementation of Zero Trust, which manufacturers should consider when implementing a Zero Trust architecture [9]. The tenets in the context of manufacturing are outlined in the following.

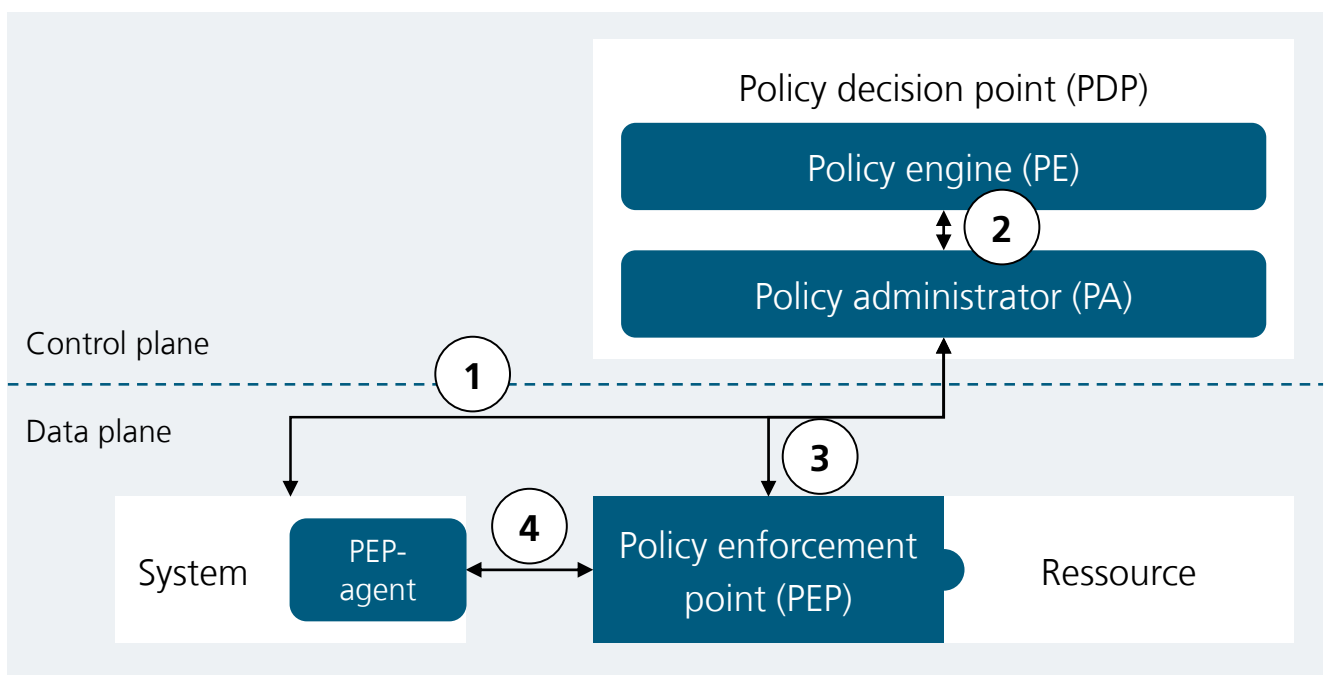


Figure 8: Zero Trust reference architecture and logical components.

**1. All data sources and computing services are regarded as resources.**

In digitalized manufacturing, production IT components such as programmable logic controllers (PLCs), sensors, edge PCs, etc. are also resources and must be considered as part of a comprehensive Zero Trust architecture. Particular attention must be paid to legacy devices that are frequently used in the production environment.

**2. All communication is secured regardless of the location of the network.**

In addition to user access, the encryption of Machine-to-Machine (M2M) communication could be required, particularly when interactions between IT- and OT systems occurs. An example could be the usage of edge cloud systems for process monitoring and predictive maintenance. For real-time cloud access to control or sensor data, both hardware and software measures are needed to minimize latency from encryption. Depending on real-time needs, encryption can be implemented using Transport Layer Security (TLS) at layer 5, Internet Protocol Security (IPSec) at layer 3, or MAC Security (MACSec) at layer 2, with lower layers typically supporting stricter real-time requirements [12].

**3. Access to individual resources is granted per session.**

In the context of digitized production environments, access to resources such as sensors, PLCs, or applications is facilitated through temporary sessions that feature customized permissions based on established policies. As a result, each new session must be verified and evaluated by a Policy Decision Point (PDP).

**4. Access to resources is determined by dynamic policies.**

Dynamic policies enable real-time adjustments to access authorizations for resources based on predefined factors. In contrast to traditional whitelisting methods, these policies can incorporate various criteria, such as geographical restrictions, time-based controls, User Behavior Analytics or hybrid approaches.

**5. Monitoring the integrity and security of all resources.**

Monitoring network and resource behavior necessitates the integration of a sensor-probe system for each resource, including an Intrusion Detection and Prevention System (IDS/IPS). In digitized OT environments, the diverse

range of vendors, along with their proprietary operating systems, applications, and network protocols, can complicate the integration of solutions like Security Information and Event Management (SIEM) or Extended Detection and Response (XDR). This complexity demands a strategic approach to achieve effective security monitoring across diverse systems, ultimately contributing to the establishment of a secure ecosystem.

**6. Authentication and authorization are dynamic and strictly enforced before access.**

In digitalized OT environments, dynamic authentication and authorization are critical to maintaining security and operational integrity. Access controls are not static; instead, they adapt based on real-time factors such as user roles, operational conditions, and context. Depending on the access requirements for a user or application, certificate-based authentication and authorization can facilitate time-based access, complete with detailed logging. A Public Key Infrastructure (PKI) based on the X.509 standard can be utilized to efficiently manage cryptographic keys and issue the required certificates for authentication and authorization.

**7. Collection of data on all resources and networks to improve security.**

Data collection aligns seamlessly with the principles of Industry 4.0 and can be leveraged within a Zero Trust framework to strengthen security measures. This includes establishing baseline processes, training AI-based anomaly detection systems, managing vulnerabilities, and analyzing access logs. These strategies collectively enhance the security posture of digitalized OT environments, enabling more effective threat detection and response.

As derived from the principles of the Zero Trust Maturity Model (ZTMM), the implementation gradient can be illustrated across five distinct pillars, allowing for incremental advancements toward optimization over time. These pillars, depicted in Figure 9, encompass identity, devices, networks, applications and data. Each pillar outlines some key aspects for the integration of Zero Trust in production environments related to three overarching functions: Visibility and Analytics, Automation and Orchestration, and Governance. These functions align with the demands of digitalized production in Industry 4.0, collectively creating a strong foundation for resilient manufacturing.

Deployment cycle for the implementation of Zero Trust in production environments

As a best practice for the integration of Zero Trust in production environments, the first step will be given by the establishment of OT Cybersecurity Governance, followed by the definition of OT specific policies and procedures. This initial step can be aligned with recognized standards such as ISO/IEC 27001 or IEC 62443-2-1. Simultaneously, the company must enhance its capabilities in automation, orchestration, visibility and analytics to facilitate the integration of selected processes into a Zero Trust architecture in accordance with Figure 8.

Once processes are established, such as through the implementation of an Information Security Management System (ISMS), the deployment cycle can be leveraged to integrate Zero Trust within a change management process as shown in Figure 10. During the preparation and categorization process, a comprehensive assessment of all resources and users, including business processes, must be carried out and inventoried. Depending on the size and complexity of the system under consideration and the company's level of maturity, this step can take a considerable amount of time. After the inventory is complete, the critical processes with the highest associated risks must be identified. A suitable process is then chosen based on the risk analysis and subsequent risk quantification. For the pilot phase, it is advisable to select non-critical processes to mitigate the impact of potential failures during the initial rollout.

If the risk assessment is completed, the implementation phase for the selected candidate process begins. Based on the Zero Trust reference architecture shown in Figure 8, the logical components, such as PDP and PEP, must be developed first. For example, the PDP in the control plane could be deployed either on-premises or in public cloud environments. The PDP and the agent in the data layer must be implemented directly on the relevant systems and resources. If legacy systems are in use or if implementing a PDP on proprietary hardware or software is too complex or costly, integrating a Zero Trust security gateway as a PEP could be considered.

Once the Zero Trust reference architecture for the selected candidate process, along with the policy configurations, is established and tested, the pilot system can be deployed in a real operational environment, a realistic test lab or a sandbox. By establishing a baseline activity pattern, policies can be refined based on practical experience. If the baseline activity is established and evaluated, companies can either expand their strategy to include new candidate processes or enhance their existing Zero Trust architecture by leveraging the five pillars of the ZTMM, as outlined in Figure 9. For instance, an IDS could be integrated for the selected candidate process. If the evaluation is found lacking, the deployment cycle can be adjusted and restarted from the beginning.

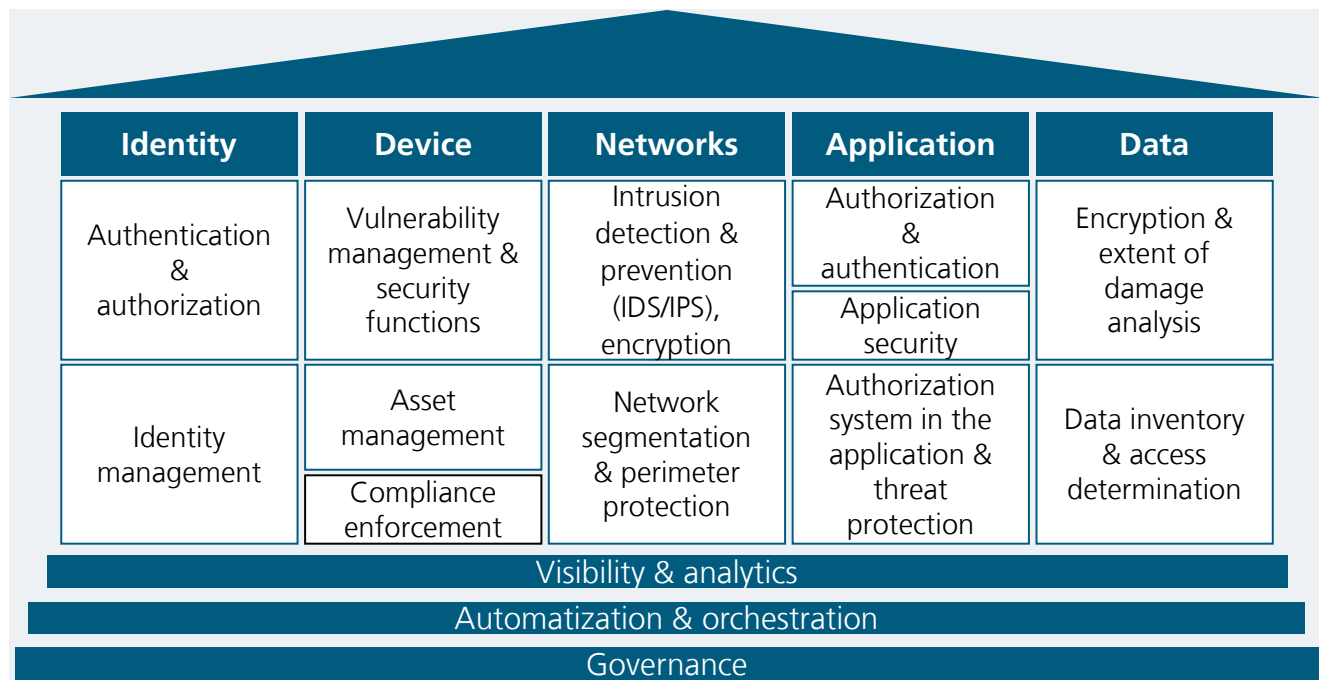


Figure 9: Pillars of the Zero Trust integration model.

In general, organizations can implement a Zero Trust architecture in production environments through various strategies. Common methods that align with the seven tenets of Zero Trust include enhanced identity governance, micro-segmentation, and software-defined perimeters [11]. Each method adheres to Zero Trust tenets but may prioritize different

components. The choice of approach typically depends on specific use cases and existing policies, with some being easier to implement than others. While alternative methods remain feasible, they may necessitate more substantial changes to current business processes.

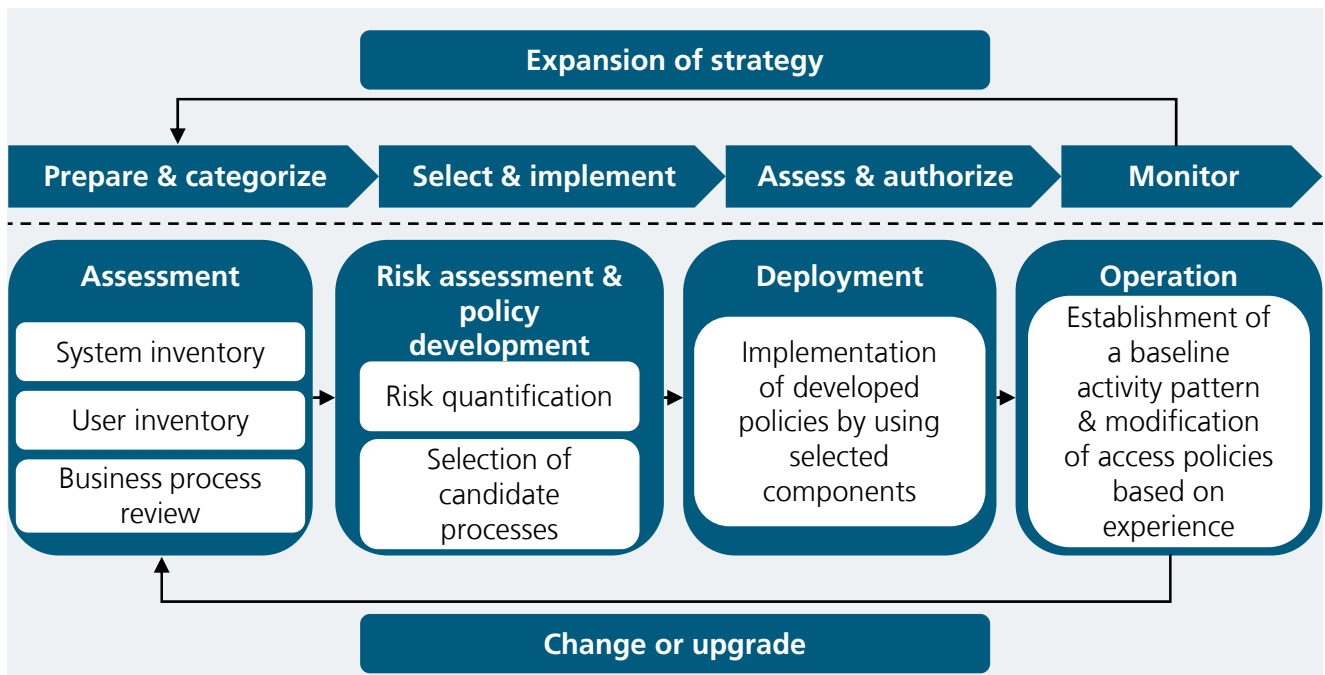


Figure 10: Hybrid Zero Trust architecture deployment cycle.



# Conclusion

The recent number of security incidents shows that cybersecurity will remain a pressing issue for organizations, driven by emerging threats and tightening security and regulatory requirements. As these challenges evolve, the importance of robust cybersecurity measures is likely to increase significantly. As a reference, expenditure on IT security in Germany will increase by 13.1 % by 2024 and will exceed the 10 billion euro mark for the first time [13]. However, there is a significant gap compared to the total loss of 266.6 billion euro forecast for Germany in 2024.

In this context, Zero Trust could provide a robust strategy to mitigate the risks associated with digitalized and connected production, enabling companies to improve their cyber resilience and protect themselves from future threats. Adopting a Zero Trust architecture in digitalized production environments presents numerous opportunities to enhance security. By following Zero Trust principles, organizations can strengthen their security posture through targeted measures that establish explicit trust for identities, devices, networks, applications, and data. This strategy considerably lowers the risk of unauthorized access and malicious activities. Furthermore, by minimizing the attack surface through restricted resource access, Zero Trust effectively reduces potential entry points for attackers. This enables companies to meet the obligations to implement security mechanisms as prescribed by NIS2 and prepare their business for the future in the context of upcoming regulations and cyber threats.

However, the integration of a Zero Trust architecture introduces unique challenges, particularly in digitalized production environments. For instance, having a comprehensive data inventory and a clear understanding of necessary data communication is crucial for organizations. Without a well-defined awareness of permissible network interactions, access requirements, and the locations of sensitive data within the

infrastructure, the risk of integration failures rises significantly. Moreover, when planning and designing a Zero Trust architecture, it's essential to consider OT-specific requirements, including the need for high availability, low latency, and the integration of legacy systems. An inadequately designed integration that fails to align with business processes can undermine the effectiveness of a Zero Trust implementation. Furthermore, the lack of standardization, particularly in OT, along with elevated costs, presents considerable obstacles to the successful deployment of Zero Trust within production environments.

Despite the various challenges associated with the implementation of Zero Trust, organizations could position themselves for the future through a strategically planned approach, thereby adapting to the evolving threat landscape. Considering that "the path to Zero Trust is an incremental process that may take years to implement" [14], it would be advisable for companies to proactively address the concept of Zero Trust at this stage and explore the potential for incorporating initial measures, driven by regulatory obligations, into their cybersecurity planning process.

# The Digital Twin Demonstrator – Bringing the concept to life

---

**André Gilerson**

Research Fellow  
Digital Infrastructures  
Fraunhofer Institute for Production Technology IPT

**Alexander Mattern**

Research Fellow  
Production Quality  
Fraunhofer Institute for Production Technology IPT

**Prof. Dr.-Ing. Robert H. Schmitt**

Member of the board of directors of Fraunhofer IPT and  
holder of the chair for Production Metrology and Quality  
Management at the WZL | RWTH Aachen University

# Introduction

With the emergence of Industry 4.0 in the 2010s, the idea of digital twins increasingly came into focus. Digital twins were introduced as a concept to create a digital representation of a physical object, machine, or system. The rapid progress of the Internet of Things (IoT) and the availability of powerful cloud computing have further driven the development of digital twins. Even though the first concept of a digital twin has been present for several years, the integration of digital twins in the manufacturing sector still needs to improve. Digital twins are heavily dependent on high-quality and reliable data. If the data used to create and update a digital twin is incomplete or incorrect, this can affect the accuracy and reliability of a digital twin. In addition, not all required data sources may always be available, especially when dealing with older machines or systems that were not designed from the beginning to be integrated into a digital twin. Another challenge is creating an accurate and detailed digital twin. It requires extensive modeling efforts to capture all relevant aspects of the real system. The more complex the real system, the more difficult it can be to include all details in the digital twin, which can lead to modeling simplifications or approximations. Further, scaling a digital twin to large systems or complex production environments can be challenging. Therefore, the scalability of digital twins is an issue that must be considered during implementation.

The Digital Twin Demonstrator – Bringing the Concept to Life study is a follow-up study on the 2023 digital twin study. The 2023 study focused on laying the foundations for digital twins in manufacturing. The Fischertechnik Learning Factory 4.0 was used as simple hardware to demonstrate the concepts and serves as the physical model for the demonstrator. It was selected by the ICNAP community during last year's study based on two main requirements. First, the physical model needed to come preassembled for ease of setup. Second, the actuators must be controlled using industrial-grade hardware to demonstrate the transferability of the digital twin to industrial use cases. In this case, the Learning Factory 4.0 uses a Siemens S7-1500 PLC. Although most of the actuators are not industrial grade, the Siemens PLC allows us to implement the demonstrator using industrial hardware, reducing the gap between the demonstrator and real-world use cases. As a result, the 2023 study implemented a digital twin utilizing the Asset Administration Shell as a data bridge and Unity and Real-virtual.io for their visualization and simulation capabilities of the digital twin. However, essential concepts for a digital twin, such as controlling the physical asset through the digital twin, were not completed in the 2023 study. This was the motivation to create a 2024 study.

In the 2024 study, we continue working on the Fischertechnik digital twin demonstrator by expanding last year's study to support new use cases and features. Building upon the foundations laid in 2023, our focus in 2024 is on completing and extending the demonstrator to explore new technological applications within the realm of Industry 4.0. The previous study successfully created a digital twin demonstrator, highlighting key benefits and addressing challenges in integrating digital twins within manufacturing environments. Chapter 4.2 summarizes its content. This year's study aims to go further by implementing a cloud-based control system, introducing a collision warning mechanism, and developing a product defect notification system.

Integrating cloud-based control allows for enhanced remote management capabilities, ensuring greater scalability, flexibility, and real-time oversight of production systems. By leveraging cloud technology, manufacturers can monitor and adjust system parameters from any location. Furthermore, cloud connectivity paves the way for integrating advanced data analytics and artificial intelligence to optimize manufacturing processes.

Another key innovation in this study is developing a collision warning system. This feature is designed to increase safety and reduce risk in the manufacturing process by using data to detect potential hazards in real time. By accurately predicting collisions, the system can provide timely alerts, ensuring that machine operators or automated systems can take preventive action, minimize damage, and maintain production flow.

Additionally, the product defect notification system represents a significant advancement in quality control. This feature enables real-time detection of product anomalies during manufacturing, ensuring defects are identified and addressed early in production. This improves product quality, reduces waste and rework, and contributes to more sustainable manufacturing practices.

This follow-up study evolves last year's demonstrator. We aim to highlight how digital twins can improve production processes, reduce downtime, enhance safety, and improve overall product quality by addressing critical use cases such as cloud-based control, collision detection, and defect notification. The findings from this extended demonstrator will provide valuable insights into the future of digital twin applications in the manufacturing industry, offering practical, scalable solutions to complex industrial challenges.

# Summary of the 2023 study

The 2023 study concentrated on implementing the backend dataflow of digital twins, aiming to showcase the advantages of open, standardized, and machine-readable digital interfaces. The central element in this setup is the standardized interface, realized through the Asset Administration Shell (AAS) deployed on a BaSyx server, as depicted in Figure 11. This interface centralizes data exchange, making it easily manageable and updateable. The AAS offers several key features:

- 1. Asset registration:** The BaSyx Server allows assets to be registered in the AAS, capturing relevant information such as identification, properties, functions, and relationships.
- 2. Asset management:** It manages assets throughout their lifecycle, including capturing changes, managing updates and versions, and tracking asset-related events and activities.
- 3. Asset provisioning:** It makes assets available for other components within the Industry 4.0 environment, enabling seamless interaction and integration.

- 4. Metadata management:** It manages metadata of the assets, including descriptions of attributes, interfaces, states, events, and other relevant information.
- 5. Security and access control:** It implements security measures to ensure confidentiality, integrity, and availability of the assets and the AAS, providing access control based on defined permissions and roles.

The standardized interface facilitates modular and independent interaction of software components, simplifying development, maintenance, and integration of new or replacement components. This modularity also allows for the decoupling of data protocols, enabling a centralized data bridge component to dynamically translate between different protocols.

This approach empowers companies to seamlessly leverage software components from various developers, as they can interact without extensive infrastructure adjustments or protocol bridging. The demonstrator illustrates this by using different database technologies tailored to data structure requirements. For static data, such as digital product

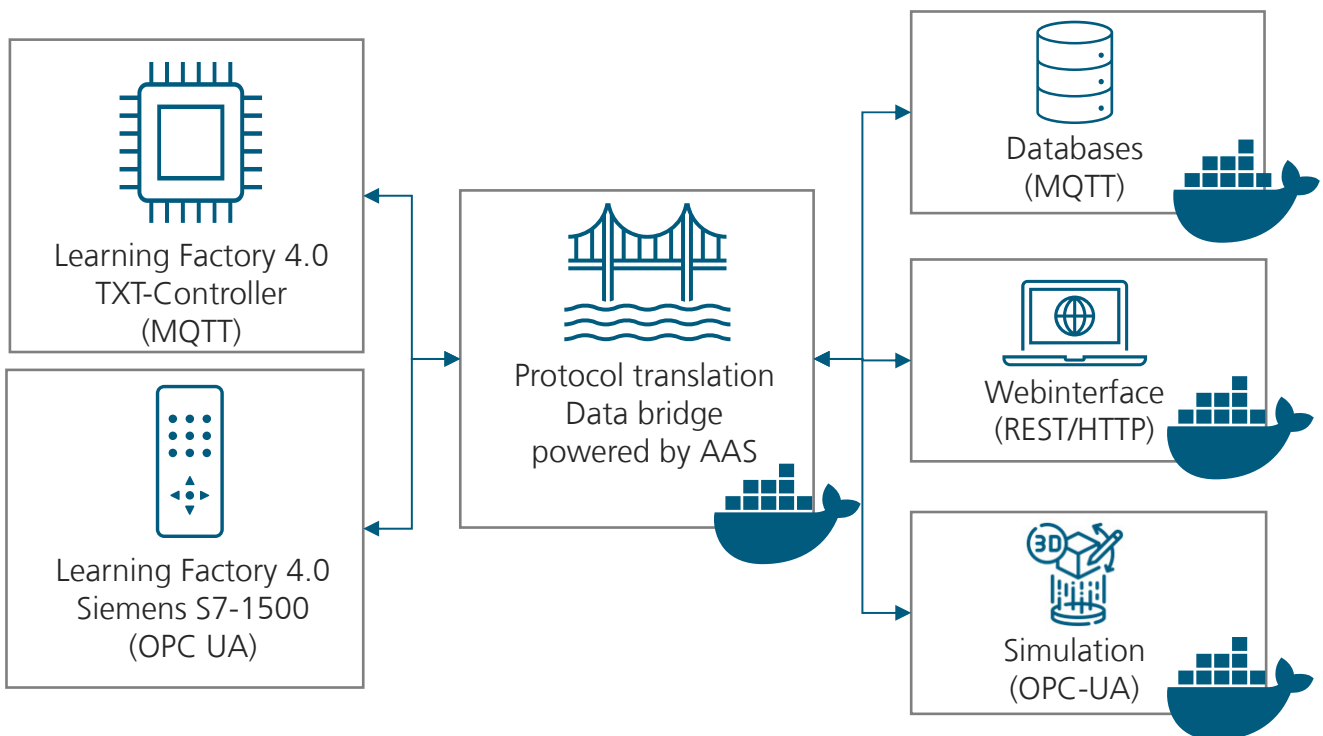


Figure 11: Data flow inside the digital twin demonstrator.

datasheets or production planning, MongoDB is employed, while dynamic, time-series-based data required for simulations is stored in InfluxDB. Both databases are hosted within their own Docker containers, with the AAS registry serving as the sole access point. The semantic description of sensor data from the Fischertechnik Learning Factory 4.0 allows the database to automatically determine the most suitable storage solution.

The results study emphasizes the benefits of a standardized and centralized data interface in creating scalable and adaptable software systems, enabling efficient and seamless integration of diverse components in the digital twin ecosystem.

## Motivating the use cases for the digital twin

During this year's study, we extended the functionality of the digital twin. Chapter 4.2.2 discusses the study goals that were planned together with the ICNAP community. Based on the goals, use cases are derived and presented in Chapter 4.2.3.

### Goals of the study

Together with the ICNAP community, we defined and prioritized the study's objectives based on their specific interests in digital twin use cases. During the meeting, four main objectives were proposed to guide the development and implementation of the digital twin demonstrator: Complete the 3D visualization of the Learning Factory 4.0 in the Unity framework; Implement a feature displaying machine errors in the digital twin, offering enhanced diagnostic capabilities; Integrate a system to inform users about product defects and improving quality control measures; Achieve full digital control of the hardware through cloud-based systems, allowing remote control and management through the digital twin.

The ICNAP community prioritized finishing the 3D visualization and enabling fully digital control through the digital twin. The third goal, which focused on product defect notifications, was acknowledged as a valuable addition but was marked as a potential option rather than a core priority for this phase. This process of goal setting ensured that the study would focus on the most relevant and impactful developments for the industry partners, with flexibility for future enhancements based on the outcomes of this phase.

In the following chapter, the three use cases are shortly presented.

### 3D visualization

The 3D visualization of manufacturing processes through digital twins provides an immersive and interactive representation of physical systems, allowing operators to visualize, analyze, and optimize processes in real time. This use case leverages advanced graphics and data integration to enhance decision-making and operational efficiency. The visualization aims to monitor manufacturing operations, including machinery status and production metrics. It seeks to enhance understanding of complex processes through visual representation the operational awareness. Lastly, it facilitates data-driven insights by integrating real-time data from the physical environment, enabling informed decision-making and rapid responses to issues.

The benefits of this 3D visualization approach are significant. By visualizing the entire manufacturing process, operators can identify inefficiencies or bottlenecks that may not be apparent through traditional monitoring methods, facilitating proactive adjustments to enhance productivity. Real time data visualization also allows for immediate insights into system performance, enabling operators to make informed decisions quickly and reducing response times to issues. Furthermore, visualizing machine operations and workflows in a 3D environment helps identify potential hazards, enhancing operator safety protocols and training. Additionally, the immersive nature of 3D visualization fosters better collaboration among team members, as stakeholders can visualize and discuss processes collectively, leading to more effective problem-solving.

### **Digital cloud-control of hardware**

The main objectives of this implementation are to provide remote access to hardware, facilitate real time data exchange, and enable proactive management of manufacturing processes. This digital control system allows operators to adjust machine settings, monitor performance metrics, and respond to issues without being physically present on-site, improving overall productivity.

One of the significant advantages of cloud-based digital control is the scalability it offers. As manufacturing needs grow, additional hardware can be integrated into the cloud system without substantial infrastructure changes. This flexibility allows manufacturers to adapt quickly to changing market demands or production requirements. Furthermore, cloud-based control enhances team collaboration by providing a centralized data-sharing and communication platform. Operators, engineers, and management can access the same information in real time, fostering a culture of transparency and informed decision-making. This collaborative environment also allows for more effective troubleshooting and problem resolution, as stakeholders can quickly identify and address issues.

### **Machine error notification**

The digital twin offers advanced visualization capabilities and can detect machine error states before they happen through simulation technology. This use case lets users test machine programs on the digital twin beforehand, without having to test them on real machines, minimizing the risk of machine failure or damage. During operation, this can also help notify the operator about previous unforeseen error states. This again shows the direct monitoring benefits that digital twins offer the operators.

### **Product defect notification**

The product defect notification system is critical to ensuring quality control within manufacturing processes. By leveraging real-time data analytics and cloud-based communication, this use case enables manufacturers to promptly identify and address product defects, thereby minimizing waste and enhancing customer satisfaction. The primary objectives of this implementation are to provide immediate alerts for product defects and enable data-driven decision-making. By establishing a robust notification system, manufacturers can swiftly react to quality issues, reducing the risk of defective products reaching the market. A significant advantage of the Product Defect Notification system is its capacity for effectively monitoring production quality by combining various data flows within the digital twin. Manufacturers can achieve a holistic view of the production process by integrating data from multiple sources – such as machine performance, environmental conditions, and material specifications. This comprehensive monitoring allows quicker identification of quality issues, as operators can correlate defects with specific machine states or ecological factors.



# Technical implementation of the digital twin

## 3D visualization and digital cloud control of hardware

The technical implementation of the 3D visualization utilizes the CAD files provided by Fischertechnik for the Learning Factory 4.0. While these CAD files contain all components with proper dimensions and placements of the Learning Factory 4.0, they offer no information about the movement capabilities of the components. Instead of creating separate animations for each movement, using regular computer graphics techniques, we implemented a custom movement system based on the established robot joint system, where each separate movement is comprised of six base movement components:

- Rotary joints: Movement around an axis
- Linear joints: Straight-line motions
- Twisting joints: Enabling rotation
- Revolute joints: Single-axis rotation
- Spherical joints: Multi-directional mobility
- Cylindrical joints: Combined movements

This allows us now to actuate each movement via virtual motors, similar to the actuation of the physical counterpart,

with the Unity Physics Engine computing the movements of the 3D models inside the digital twin. This approach not only facilitates easier development of the digital twin but also improves the realism of the digital twin by integrating complex motion dynamics simulations without having to incorporate them into 3D animation frameworks. One minor disadvantage of this approach is the visualization's reduced smoothness since all components' positions are recalculated at each time step without blending in between. However, this was regarded as inconsequential as the simulation can still run above 30 fps on a regular desktop PC without high-end graphics hardware, providing the illusion of smooth movement for most components. An example of the UI is shown in Figure 12.

A critical aspect of this implementation is transitioning from a sequential hard-coded PLC code provided by Fischertechnik to a modular code. In the traditional approach, the PLC executes a fixed sequence of operations, which limits flexibility and adaptability. By moving to a modular design, each movement is encapsulated within a separate functional block. This modular approach allows these blocks to be triggered independently through the digital twin interface.

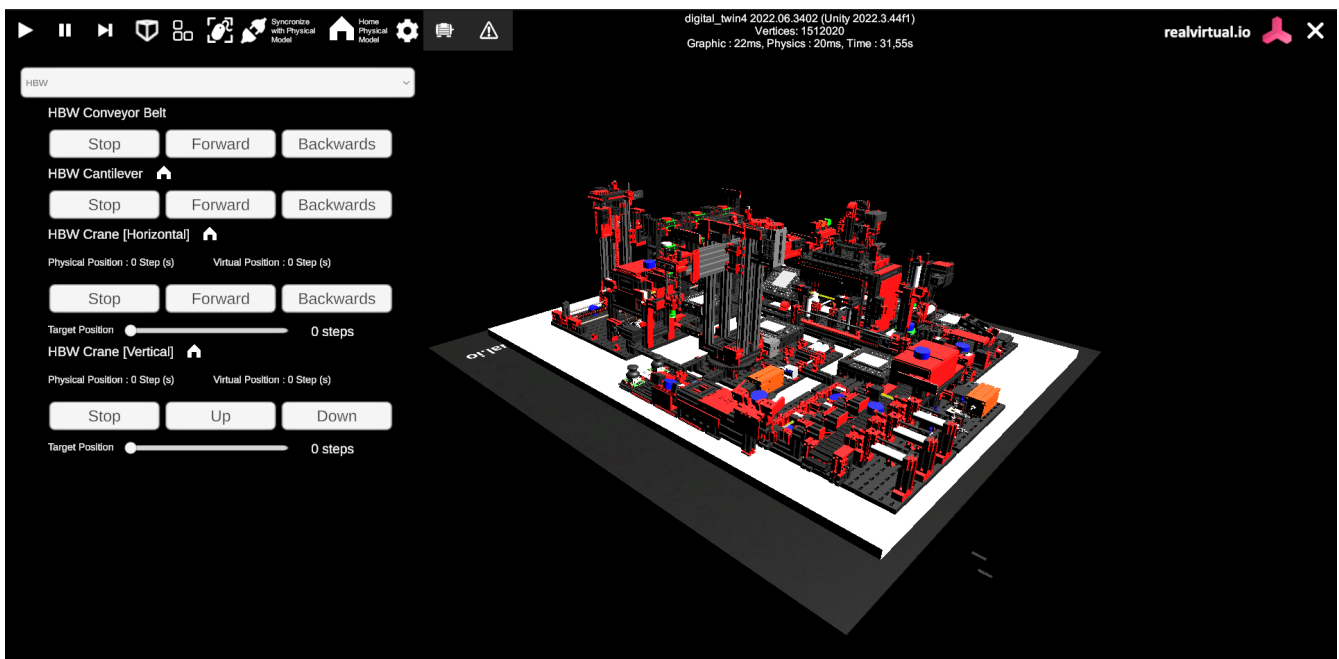


Figure 12: User interface of the digital twin of the Fischertechnik hardware.

The benefits of this modular structure are substantial. Firstly, it enhances flexibility, enabling operators to adjust individual movements without rewriting the entire control program. This adaptability is particularly useful in dynamic manufacturing environments where processes must be altered frequently. Secondly, it simplifies debugging and maintenance, as developers can isolate issues within specific functional blocks rather than navigating through a complex, interconnected sequence. Lastly, this approach promotes reusability, allowing the same movement blocks to be employed in different processes or setups, ultimately speeding up the development cycle. However, limitations are associated with the S7-1500 PLC used in the Learning Factory 4.0, particularly concerning its OPC UA server. The minimal update time for the OPC UA subscription mechanism is set at 0.5 seconds, which introduces significant latency into the system. This delay can hinder real-time responsiveness, making achieving the desired interactivity and fluidity in the 3D simulation challenging. We propose moving from the subscription mechanism to a busy polling system to address this issue. In this configuration, the Unity application will actively poll the appropriate OPC UA variables for each movement at a much higher frequency. By employing separate threads for each movement, the system can independently query the PLC for updates without introducing excessive latency. This design allows for more immediate feedback in the 3D simulation, enhancing the overall user experience and ensuring operators can interact with machinery in real time.

## Machine error notification

The machine error notification system is a crucial component of the digital twin that enhances the safety and reliability of manufacturing processes. This system is demonstrated by predicting potential collisions among components in the Learning Factory 4.0 setup.

The implementation begins with creating custom collision boxes for each component within Learning Factory 4.0. These collision boxes are designed to encapsulate various physical dimensions and operational boundaries, such as robotic arms, conveyor belts, and other machinery. By accurately modeling these components, we ensure the collision detection system can effectively simulate real-world interactions. Once the collision boxes are established, the Unity Physics Engine handles

real time collision detection. The engine continuously monitors the positions and movements of the components, checking for intersections between the collision boxes. When a potential collision is detected, the system triggers predefined responses, including notifications to operators, alerts to halt operations, or visual cues in the 3D simulation indicating the impending collision.

Implementing a collision detection system offers several potential benefits for actual industrial processes.

- 1. Enhanced safety:** By predicting collisions before they occur, manufacturers can significantly reduce the risk of accidents and injuries in the workplace. This proactive approach to safety creates a more secure environment for operators and machinery alike.
- 2. Reduced downtime:** Early detection of potential collisions allows for timely interventions, preventing equipment damage and subsequent downtime. In a manufacturing context, unplanned stoppages can be costly in terms of lost productivity and repair expenses. By minimizing the likelihood of collisions, manufacturers can maintain smoother operations.
- 3. Improved efficiency:** With collision detection systems in place, operators can optimize their workflows without fear of unintended interactions between components. This increased confidence can lead to more efficient use of machinery and resources, ultimately driving higher productivity.
- 4. Training and simulation:** The collision detection mechanism can also serve as a valuable training tool for new operators. Trainees can gain hands-on experience in managing machinery safely and effectively by simulating potential collision scenarios within the digital twin. This educational aspect enhances skill development while reducing the risk of real-world incidents.

The collision detection system seamlessly integrates into the digital twin framework, allowing real-time monitoring and control. Operators can visualize the current state of the manufacturing process, including any detected collisions, within the 3D simulation. This integration provides a comprehensive view of operations, enabling informed decision-making and quick responses to potential issues. The machine error notification system is designed to be highly customizable, allowing manufacturers to tailor the parameters and thresholds for collision detection based on their specific operational needs. This flexibility ensures that the system remains relevant and practical across various manufacturing.

## Product defect notification

An initial challenge we faced was that the Learning Factory 4.0 does not come equipped with sensors that could be directly utilized for detecting product errors. To address this limitation, we adopted the approach of creating simulated virtual sensors. These virtual sensors are programmed to simulate the behavior and data output of real-world sensors that would typically monitor various aspects of the manufacturing process, such as dimensional accuracy, surface quality, and assembly precision. The virtual sensors generate data based on predefined error conditions and scenarios that mimic potential real-world defects.

The AAS framework provides a standardized, centralized interface for managing and exchanging data among various components of the digital twin ecosystem. In this setup, the simulated virtual sensors generate data captured and registered within the AAS. The AAS framework enables real-time

monitoring and management of this data by storing detailed metadata about each virtual sensor, including identification, properties, and error conditions. When an error is detected, the AAS triggers an immediate notification, leveraging its robust security and access control mechanisms to alert the right personnel promptly. This integration ensures that all data exchange and error notifications are handled in a modular, scalable, and secure manner, facilitating efficient interoperability among the digital twin components and enhancing the overall reliability of the manufacturing process.

## Conclusion

In conclusion, this study builds upon the foundations established in the previous year's digital twin demonstrator study, significantly expanding its capabilities to address critical Industry 4.0 use cases. By integrating cloud-based control, a collision warning mechanism, and a product defect notification system, the digital twin is now more versatile and relevant to modern manufacturing environments. Implementing 3D visualization through advanced modeling and data integration enhances operational efficiency by allowing real-time monitoring and optimization of production processes.

The introduction of cloud-based control adds flexibility, scalability, and real-time data exchange, enabling remote management of manufacturing systems. This promotes enhanced collaboration and faster response times. The collision warning system improves safety and operational flow by detecting potential hazards before they occur, while the product defect notification system supports quality control through early detection of anomalies, reducing waste and rework.

Through these innovations, the study not only extends the functionality of the Fischertechnik demonstrator but also provides valuable insights into the future of digital twin applications. The findings illustrate how digital twins can improve productivity, enhance safety, reduce downtime, and ensure better product quality, offering practical solutions to meet the evolving needs of the manufacturing industry.

# Seamless AI integration through Plug & Produce approach

---

**Jan Hendrik Hellmich**

Research Fellow

Production Quality

Fraunhofer Institute for Production Technology IPT

**Ines Groß-Weege**

Research Fellow

Adaptive Production Control

Fraunhofer Institute for Production Technology IPT

**Prof. Dr.-Ing. Robert H. Schmitt**

Member of the board of directors of Fraunhofer IPT

and holder of the chair for Production Metrology and Quality  
Management at the WZL | RWTH Aachen University

# Introduction

## Motivation

In recent years, artificial intelligence (AI) applications are emerging in the field of manufacturing industry enabled by the greater availability of data. The rise of Industry 4.0 technologies has driven significant advancements in the realm of AI, fostering a greater adoption of this innovative technology. As a result, manufacturing environments are becoming increasingly dynamic and connected envisaging higher agility, productivity, and sustainability. Solutions such as data-driven predictive analytics and assisted decision-making are thereby contributing to an improved product quality, increased performance and cost reductions [15], [16].

However, the industrial adoption of such solutions is characterized by a high degree of complexity. Various challenges are faced such as real-time processing needs, high security requirements, the integration of heterogeneous devices as well as large data volumes to be handled. Numerous organizations are not adequately prepared to address these issues which impedes the integration of AI solutions in the industrial context [15]. In connection with this subject, the vision of “Plug & Produce” was addressed. It is defined as the “capability of a production system to automatically identify a new or modified component and to integrate it correctly into the running production process without manual efforts and changes within the design or implementation of the remaining production system” [17]. This means, machines and devices can be interconnected instantaneously without the need of a specific driver installation or change in setting configurations. New components can be added rapidly and seamlessly, thus increasing flexibility, efficiency and interoperability in the manufacturing environment [18], [19].

With regard to the integration of AI applications to the shop floor, this concept is usually not applied as current AI pipelines are nowadays often detached from the production environment. They mostly only receive manual data input (e.g. csv files) and conversely, do not automatically deliver the AI output to the physical system. That is why the question arises how it can be achieved to integrate AI applications seamlessly to the shop floor by using a Plug & Produce approach [17], [18].

For the connection between AI application and physical assets, an Industrial Internet of Things (IIoT) connectivity framework has to be established. It incorporates the communication beginning from the physical layer (e.g. Ethernet or WiFi) over the transport layer which includes different machine protocols to the distributed data interoperability and management layer [19].

However, the challenges can be summarized as follows:

- The lack of clarity regarding the specific technologies required to link AI applications and machines poses a challenge [20].
- The market offers a diverse array of technologies, each with distinct characteristics [21].
- Selecting the appropriate option tailored to individual needs demands considerable expertise and research, thereby complicating the decision-making process [22].
- In addition, a structured approach for this problem is often not defined due to the individual requirements and different implementation options [23].

The ICNAP study “Seamless AI Integration through Plug & Produce approach” aims to address these challenges. Three primary objectives have been identified which form the foundation of this guideline:

- A comprehensive overview of available technologies on the market will be created, enabling customers to quickly identify the most suitable option for their specific use case.
- Key criteria will be established to facilitate the selection of suitable technologies and accelerate the conceptual process.
- A structured procedure will be derived, which should guide companies on how to seamlessly integrate AI into their production environments.
- Overall, the guideline envisions to reduce the effort required for technology research, enabling customers to transition into the implementation phase with greater efficiency and speed.

## Structure of the report

The study aims to address these gaps and challenges and provides a guideline for companies to realize a seamless AI integration to the physical asset.

After a comprehensive description of the initial situation and the general concept of the study in the following chapter 5.2, the next chapter 5.3 presents the technologies relevant for connecting and linking AI applications to the machines. This chapter forms the scientific basis for identifying evaluation criteria for selecting the right technologies for the respective use case.

This is followed by a comprehensive analysis of the technologies, such as interfaces, databases and IIoT platforms, which play an important role in the development and implementation of seamless AI integration.

Chapter 5.5 presents the developed procedure of the guideline, which is made up of the previously developed content. The core of the guideline is a questionnaire, which is intended to support the company in the selection of technologies. A demonstrator will illustrate the relevance and selection of technologies for practical implementation in chapter 5.6. Finally, the results are summarized and a brief outlook for further work is given.

## General concept of the study

The overall goal of the study is to support the identification and selection phase of suitable technologies for realizing a seamless AI integration at the shopfloor. An attempt is being made to achieve this goal by developing a comprehensive guideline to support the selection process for the key technologies.

The overview shown in Figure 13 serves to clarify the selected technology types which were defined as relevant for AI integration at the shopfloor. The aim is to realize a suitable connection between physical machines and AI applications into the existing environment as easily as possible.

But for most of the companies, it is not yet clear which technologies are available on the market and which are best suited to their own use case. To close this gap, this guide dives deeper into the selected technologies and analyzes which technologies meet which criteria that are relevant to the integration process and depend on different use cases.

Furthermore, in most cases, there is already a machine that performs a specific production step and is equipped with specific data interfaces. On the one hand, commands can be transmitted to the machine. On the other hand, it is also possible to retrieve production data from the machine to generate information from the operation.

In order to send the data between the machine interface and the AI application, different technologies like databases, middleware, machine protocols or IIoT platforms can be used to transfer the data and information between the two parties.

- A. A middleware system is often required to distribute the data further between the devices. This makes it possible, for example, for several AI applications to receive production data from one machine. Furthermore, predictions made by the AI application can be transmitted back to the machine with any degree of automation.
- B. For saving important data, it makes sense to store it in form of a database that is connected to the middleware. This also allows the AI application to incorporate historical data into the prediction.
- C. If an IIoT platform is already in use or is to be used, this software platform already includes some components such as databases, middleware or various interfaces for connecting the various machines and devices in the network. Integrated AI applications of the IIoT platform can also be used or an external application can be connected to the platform.
- D. Furthermore, the selection of the correct machine protocol is crucial for the transfer of information. Factors such as data speed, information size or usability of the protocol can be decisive for the selection of the protocol and success of the complete use case.

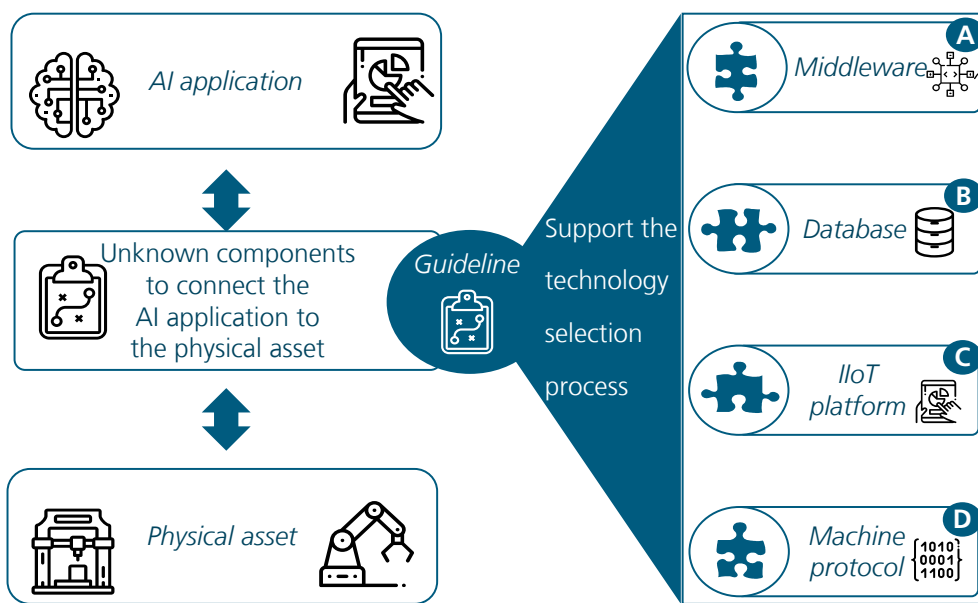


Figure 13: General concept of the study—Integration of AI application through Plug & Produce.

## Theoretical background

In the scope of this study, several technologies in the context of IIoT will be presented and evaluated. This chapter aims at giving a brief overview about relevant technologies that need to be considered in the seamless AI integration process through Plug & Produce.

Furthermore, criteria have been elaborated that have an influence on the decision as to which AI integration strategy is best suited for the respective use cases. These criteria form the basis of this study on which the further procedure is built. Consequently, this chapter dives deeper into the selected criteria identified as crucial for the seamless AI integration process.

### Relevant technologies along the seamless AI integration

After the completion of an extensive scientific research, core technologies along the seamless AI integration were identified in the field of industrial applications. These are middleware architectures, databases, IIoT platforms, communication protocols and AI applications. This section provides a short introduction on those core components building the connection between AI application and machines as depicted in the study's big picture.

#### Middleware

Middleware is a software layer that acts as an intermediary between different applications or components in a distributed system. It supports various protocols, such as HTTP, gRPC, or WebSocket, enabling seamless communication and data exchange. In the context of AI integration, middleware provides a standardized interface for deploying models, managing data flow, and orchestrating processes across multiple platforms. It ensures that AI components can interoperate efficiently with existing software infrastructures, regardless of underlying technology differences. This facilitates real-time data processing, model execution, and decision-making. By abstracting complex communication tasks, middleware enhances the scalability, reliability, and performance of AI-driven systems [24]–[26].

#### Databases

Databases are structured collections of data that are stored and accessed electronically. They are crucial for storing, managing, and retrieving large amounts of information efficiently.

A distinction is made between relational, NoSQL, object-oriented and hierarchical databases. While relational databases

save data in a tabular format and are therefore only suited for structured data, NoSQL databases can also handle unstructured and semi-structured data. Subcategories of this type of databases are key-value, document and graph databases for example. In object-oriented databases, data is saved in form of objects containing both attributes and methods. Hierarchical databases organize data in a tree structure, where each child node has one parent node and parent nodes can have multiple child nodes [27]–[30].

### **IIoT platforms**

Industrial Internet of Things (IIoT) platforms are specialized frameworks that facilitate the connection, management, and analysis of industrial devices and systems. These platforms enable the integration of various industrial protocols, data storage solutions, and analytics tools to optimize industrial operations [31].

Various tech companies offer their own IIoT platform such as Amazon, Microsoft or Siemens [32]–[34].

More than 620 IIoT platforms exist by now and the market is steadily growing [31].

### **Communication protocols**

Machine protocols are standardized communication methods used to enable data exchange between different machines, devices, and systems. These protocols ensure that machines can interact seamlessly, even if they are from different manufacturers.

Examples for well-known protocols are MQTT, OPC UA, AMQP, CoAP and MTconnect. In the following these protocols are presented in more depth.

MQTT, which stands for Message Queuing Telemetry Transport, is a lightweight and reliable protocol designed for message transmission across networks in a one-to-many distribution model. Its efficiency makes MQTT one of the leading choices among publish/subscribe communication solutions [35], [36].

Open Platform Communication Unified Architecture (OPC UA) is designed to resolve the issue of interoperability among hardware devices by offering a standardized communication framework. Created by the OPC Foundation specifically for industrial automation, it aims to merge all existing protocols into a single, cohesive data model. It is also not just a protocol, but a meta modeling language, which serves the semantic modeling of information [37]–[39].

The Constrained Application Protocol (CoAP) is designed for communication between Internet of Things (IoT) devices that have limited resources. Thus, it is ideal for low-power devices or devices with small space of memory and narrowband networks with poor connection quality [35], [38].

Advanced Message Queuing Protocol (AMQP) is an open-standard protocol that allows for asynchronous communication by enabling messages to be stored in a queue. AMQP is designed to ensure security, reliability, and seamless interaction with other systems [35], [38].

MTConnect is an open, non-proprietary, and extensible standard that leverages XML to facilitate enhanced interoperability among machines. As a one-directional read-only protocol, MTConnect is typically employed for machine monitoring purposes [39], [40].

### **AI application**

In production, AI applications such as predictive quality, predictive maintenance, quality control, and anomaly detection can offer significant benefits for companies. Predictive quality enables the forecasting of product quality through the analysis of historical data, allowing for early detection of potential quality issues. Predictive Maintenance aims to predict maintenance needs and possible machine failures, minimizing unplanned downtimes and enhancing efficiency. Quality control utilizes automated monitoring and analysis of production processes to ensure compliance with quality standards, thereby guaranteeing product consistency. Finally, anomaly detection helps to identify unusual patterns or deviations in production data, enabling quick resolution of potential problems [41].

### **Criteria to evaluate core technologies**

In this study, 18 different criteria were defined that need to be kept in mind when integrating AI into the shop floor. They were selected after screening various research papers about the integration of IoT technologies in industrial environments and interviewing experts in the field. In order to focus on the decisive technologies, the various criteria are divided into different categories and will be matched to them individually. Furthermore, the complete criteria list is shown in Figure 14.

By looking from various perspectives, different criteria get into focus. From an AI perspective for example, it is crucial to ascertain whether the application must operate in real-time, or if the decision process should be automated or not. Meanwhile, from a machine perspective, it is essential to understand the



available device resources. Conversely, the connection perspective offers insights into the significance of data security, for instance. These example criteria are supporting the selection process to find tailored technologies for the individual use case of a company and gives the developer a better sense of which properties of his solution he needs to pay attention to.

Furthermore, it is important to define each criterion once in order to create a uniform understanding. In the following a definition of one criterion is provided exemplarily for each of the three perspectives:

- **Real time capability:** The near-instantaneous processing of data and execution of actions so that systems can respond and adapt to changes and inputs as they occur.
- **Horizontal scalability:** The ability to manage an increasing number of devices.
- **Reliability:** A reliable protocol ensures the successful delivery of data to the intended recipients, where the data is received in full, without errors and in the correct order.

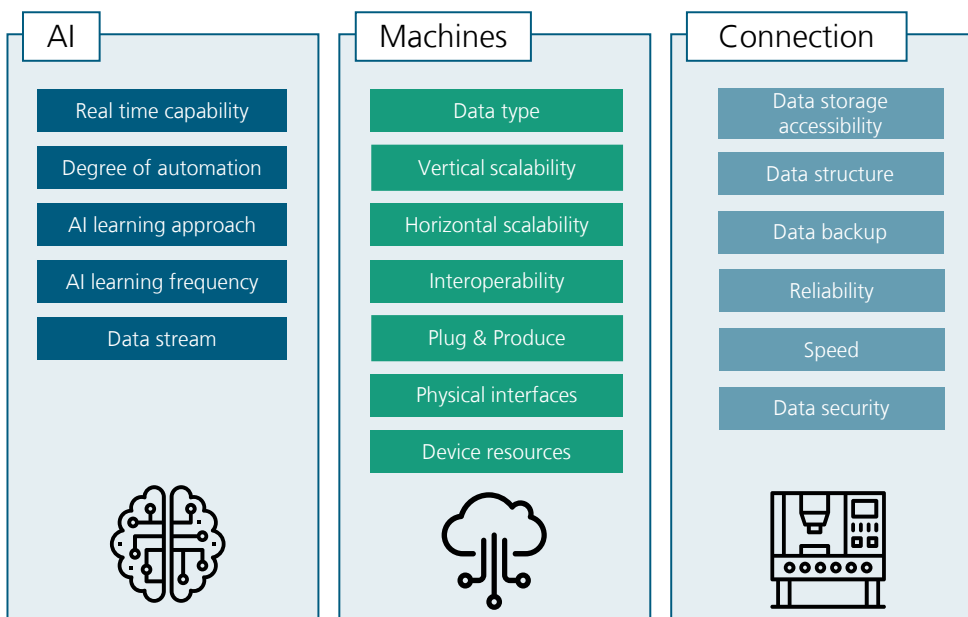


Figure 14: Selected criteria for technology evaluation.

# Evaluation of selected technologies

In this chapter, the core technologies presented in section “theoretical background” are evaluated regarding the elaborated criteria introduced in its subsequent section. The aim here is to provide a comparable overview of individual technologies, which are evaluated on the basis of selected criteria. Hence, it builds the foundation on identifying which technologies offer the highest suitability for the connection components between machines and AI application, depending on the use case.

The evaluation of the core components was done by a comparison of the technologies’ properties to the selected criteria using Harvey balls. These give an indication of the extent to which a criterion is fulfilled. The following applies: 0 % means the criterion is not fulfilled at all, 100 % that the criterion is very well or optimally fulfilled.

## Machine protocols

In April 2024, an ICNAP workshop was conducted for the 24 member companies of the ICNAP community. The community was asked to specify which industrial technologies they are

using in their production processes. Regarding machine protocols, the community stated, that they deployed MTconnect, MQTT and OPC UA among others.

After evaluating the results of the ICNAP workshop and screening various research articles regarding relevant communication protocols in the IIoT sector, five different machine protocols have been selected for further evaluation. These are MQTT, OPC UA, CoAP, AMQP and MTconnect.

The five machine protocols were evaluated regarding nine different criteria. These are data type, vertical and horizontal scalability, interoperability, device resources, reliability, speed and data security.

After having elaborated the protocol’s properties regarding every single criterion, they were transferred in a graphical representation using the above-described Harvey balls. The results are shown in Figure 15.

	MQTT	OPC UA	CoAP	AMQP	MTconnect
Data type	Binary, XML, JSON	OPC Binary, OPC XML, OPC JSON	Binary, SenML, JSON, CBOR, XML, EXI	Binary, XML, JSON, own type system	XML
Vertical scalability		*		*	
Horizontal scalability					
Interoperability					
Device resources					
Reliability					
Speed					
Data security					

\* Dependent on implementation

Figure 15: Evaluation between selected criteria and interfaces [38], [39], [42]–[45].

As can be seen, every evaluated protocol except MTconnect optimally fulfills the “Horizontal Scalability” criterion meaning that an increasing number of devices can be connected [42].

Regarding the criterion “Device Resources”, the Harvey balls were filled if the protocol has low requirements on the device resources and left empty if high requirements are needed. Since MQTT and CoAP are especially suited for communication between devices with limited resources as low-power devices or narrowband networks with limited connection quality, they optimally fulfill the criterion [35], [43].

The reliability and speed criteria were thereby dependent on the underlying network communication protocols UDP and TCP. Those communication protocols relying on UDP as OPC UA and CoAP have their strengths in a fast data transfer rate. Whereas MQTT, AMQP and MTconnect do not focus on speed but on transferring every message reliably by guaranteeing message receipt, the correct message order and preventing duplication. These are features that are enabled by the TCP protocol [42], [44].

## Databases

In the scope of this study four different categories of databases presented in section 5.3.1 were examined: Relational, NoSQL, object-oriented and hierarchical databases. Regarding the NoSQL database three subcategories were analyzed in more detail namely key-value, document and graph database. An overview about the chosen databases is given in Figure 16.

The chosen criteria for the evaluation were vertical and horizontal scalability, Plug & Produce, data structure, data backup and speed.

All databases fulfill the criterion “Vertical Scalability” effectively, handling high data volumes and complex scenarios. Relational, key-value, document, and graph databases are optimized for large-scale data, while object-oriented databases handle complex data models well. Hierarchical databases, while effective for hierarchical relationships, are less suited for large data volumes.

Most databases meet the criterion “Horizontal Scalability” by easily expanding through additional nodes or servers. Relational and hierarchical databases generally scale better vertically with fewer machines.

In terms of the criterion “Plug & Produce”, integration ease varies. Relational and hierarchical databases, with their structured data and fixed schemas, require medium to complex integration efforts. Key-value and document databases offer easier integration due to their flexible schemas, whereas graph and object-oriented databases, with their complex data models, demand greater integration effort.

For the criterion “Data Structure”, relational databases use structured, tabular formats. Key-value and document databases handle data in unstructured formats like key-value pairs and JSON documents. Graph databases utilize structures for complex relationships, while object-oriented and hierarchical databases use structured formats such as objects and tree structures.

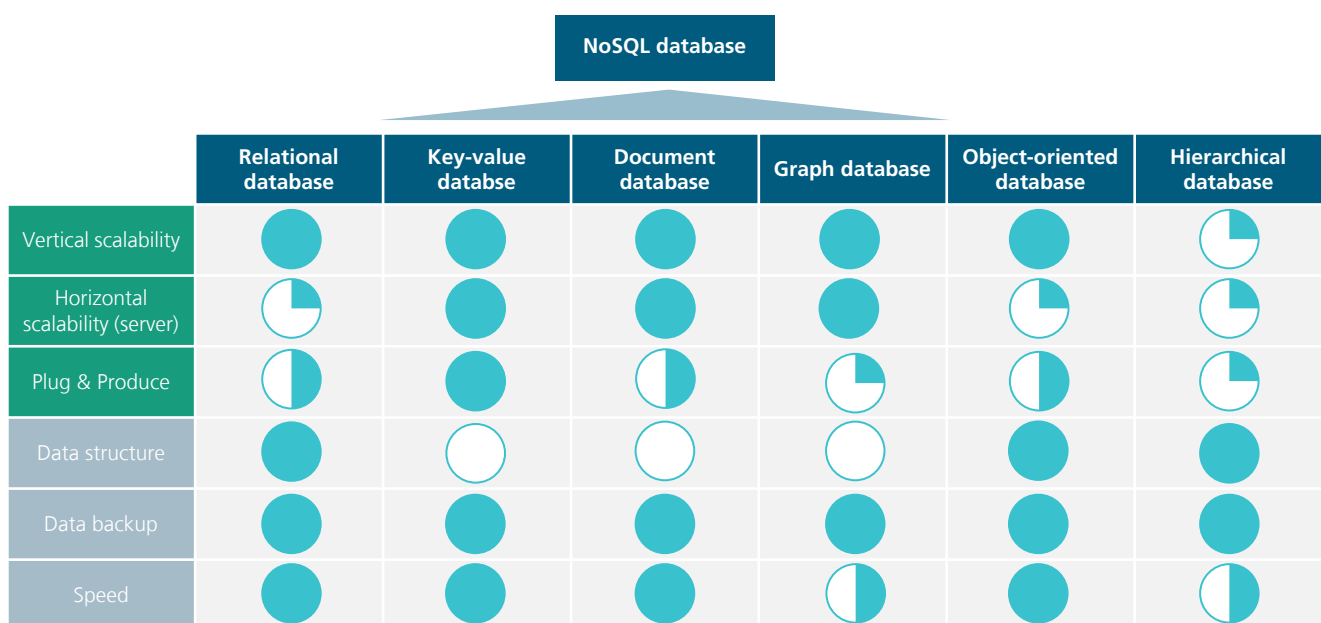


Figure 16: Evaluation between selected criteria and databases [46]–[52].

All databases fulfill the criterion “Data Backup” with well-established mechanisms for reliable data protection and recovery. Therefore, all Harvey balls are filled fully.

Regarding the criterion “Speed”, most databases provide high performance with fast queries and transactions. Hierarchical and graph databases focus more on managing complex queries and data structures, with speed being a secondary concern

IloT platforms

In this study, seven of the most widely recognized IloT platforms were chosen and evaluated considering ten of the defined criteria. These platforms are Amazon AWS IoT, Bosch IoT Suite, GE Predix, IBM Watson IoT Suite, SAP Leonardo, Siemens Mindsphere and Microsoft Azure IoT Suite which were selected after extensive scientific research about the most known platforms in literature [53].

The ten chosen criteria are real-time, data stream, protocols, integrated AI application, openness/ external AI integration, vertical and horizontal scalability, interoperability, data security and data storage.

All of the evaluated IloT platforms can perform in real-time or near real-time. They are all both vertically and horizontally scalable meaning that the data volume as well as the number of devices can be scaled according to the customer’s needs. Since every IloT platform reviewed is running on the cloud, the data storage criterion is optimally fulfilled too. Furthermore, data security methods are included in all the IloT platforms evaluated [31], [53].

One difference between the platforms lies in the fact that not every platform allows integrating own AI applications. For example, Amazon AWS IoT and GE Predix support the unrestricted integration of customers’ own applications while IBM Watson IoT Suite and SAP Leonardo do not. The number and type of supported machine protocols as well as the capability of interoperability also varies greatly [31], [53].

MQTT is supported by all of the evaluated platforms, OPC UA by most of them. CoAP is only usable within SAP Leonardo and Siemens Mindsphere [53], [54].

These points are crucial to consider in selecting the optimal IloT platform.

There does not exist a single IloT platform from the ones evaluated that satisfies all ten criteria. As with other IloT technologies presented, it is therefore important to weigh up which criteria are particularly important, and which are less relevant for the own use case in order to arrive at a well-founded decision.

AI application

For the study different typical AI use cases were selected for the guideline. These are from the areas of predictive quality, predictive maintenance, quality control and anomaly detection. In addition to these use cases, there are others such as monitoring & diagnostics, layout optimization, ramp-up optimization and more, which are not considered in detail due to the scope of the study [41]. One typical example scenario for the selected use case was defined and will be described below. The example serves to illustrate the practical use of the guideline.

Selected use case: Visual, AI-based quality control for recognizing and adjusting the order of letters

In this use case, an AI application is utilized for image-based quality control. A camera captures images of various letters that need to be arranged in the correct order by individual movers. If the letters are not in the correct order, the AI communicates this information and sends the necessary adjustments to the machine. The movers then change the positions of the letters within seconds to match the new specified order. The AI application subsequently checks again to ensure that the correct formatting is in place. If the formatting is accurate, the use case concludes, allowing for the analysis of new letters to begin.

Based on the above-described example and the selected AI criteria the predictive quality use case was evaluated as presented in table 1.

Use case: visual, AI-based quality control	
Real time capability	yes
Degree of automation	Level 4
AI learning approach	Supervised learning
AI learning frequency	Discrete
Data stream	Continuous
Data type	Image data

Table 1: Evaluation of AI application: example predictive quality use case.

# Development of a procedure including a questionnaire for identification of suitable technologies

A universal solution for the seamless integration of AI applications does not exist as it is highly dependent on the specific industrial use case. Therefore, it is essential to first identify the crucial aspects that must be considered during the integration process. Armed with these insights, a target-oriented decision in selecting the best suitable technologies along a smart factory can be made more easily thereby facilitating the discovery of the optimal AI integration pipeline.

For this purpose, four steps have been developed in this guideline, which are depicted in Figure 17. As initial situation, it is assumed that the machines to be connected to the AI application are known as well as the requirements for data transmission. The four steps start with the clarification of the specific use case by selecting the use case which should be considered for the AI integration process.

Taking the criteria presented in section “theoretical background” as a foundation, a question and answer catalog has been derived which aims at facilitating the identification of important criteria for the user. An extract is shown in Table 2 which should be filled out in the second step .

Using the resulting answers of the questionnaire, the technical realization follows in the last two steps. It begins with selecting the IIoT technologies presented in chapter “evaluation of selected technologies” for the AI integration pipeline by reviewing which of them are fulfilling the as important identified criteria. Conclusions can be drawn from this as to

which databases, machine protocols and IIoT platforms are best suited for the use case. The last step consists of the actual implementation of the connection.

The procedure for the implementation starts with configuring the machine protocol via communication parameters as the IP address and the port. After that, data points must be defined (e.g. sensors, actuators) that are to be monitored and transmitted.

Once this has been done, there remain two options to choose from. The option to connect the developed AI application via a database to the shop floor. Or the option to use an IIoT platform and benefit from the additional feature of deploying already integrated AI tools. Since all IIoT platforms already incorporate data storage capabilities, an additional database is not needed.

For the first option (“Connection via database”) the connection between the chosen database and the machines has to be set up via the selected machine protocol. Afterwards the AI application has to be connected to the database.

The second option (“Use of IIoT platform”) includes registering all machines on the IIoT platform as devices. In the following the data transmission from the machines to the IIoT platform has to be configured via the chosen machine protocol. After that the AI application can be connected to the platform or the integrated solution can be used.

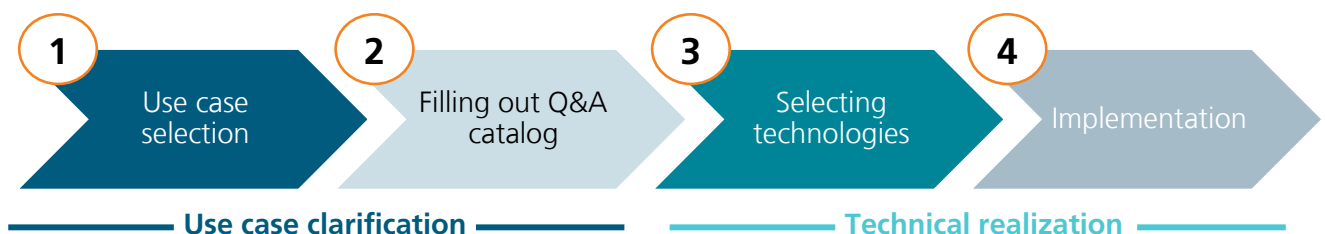


Figure 17: Guideline procedure.

Criteria	No.	Questions and Answers
<b>Data Type</b>	Q1	In what data type do your machines output their data?
	A1	Binary      Text files      Images      Audio/Video
<b>Vertical Scalability</b>	Q2	Is it important that a high data volume can be transferred per time (e.g. several megabytes per second)?
	A2	Yes                      No
<b>Horizontal Scalability</b>	Q3	Is it important that an increasing number of devices can be connected (e.g. thousands of devices)?
	A3	Yes                      No

Table 2: Extract from the created question & answer catalog.

## Use case development

To validate the guideline the presented use case from chapter AI application was selected.

The use case includes a machine moving products to specific places and involving an integrated camera system. The complete information and communication flow of the selected use case is shown in Figure 18. The products are automatically moved to the camera of a tablet which takes pictures of them (A). The tablet incorporates the AI application which after screening the images, publishes its predictions to a broker

(B-D). The machine represents the subscriber in this case and receives the predictions via the broker (D-F). Dependent on the obtained data, the machine's products are moved to a specific location managed by a control algorithm (F).

For this use case the guideline was used to choose the most suitable technologies. After the selection of the use case which represents step 1, in the next step of the guideline the question and answer catalog was filled out. Thus, the criteria reliability, interoperability and device resources were identified

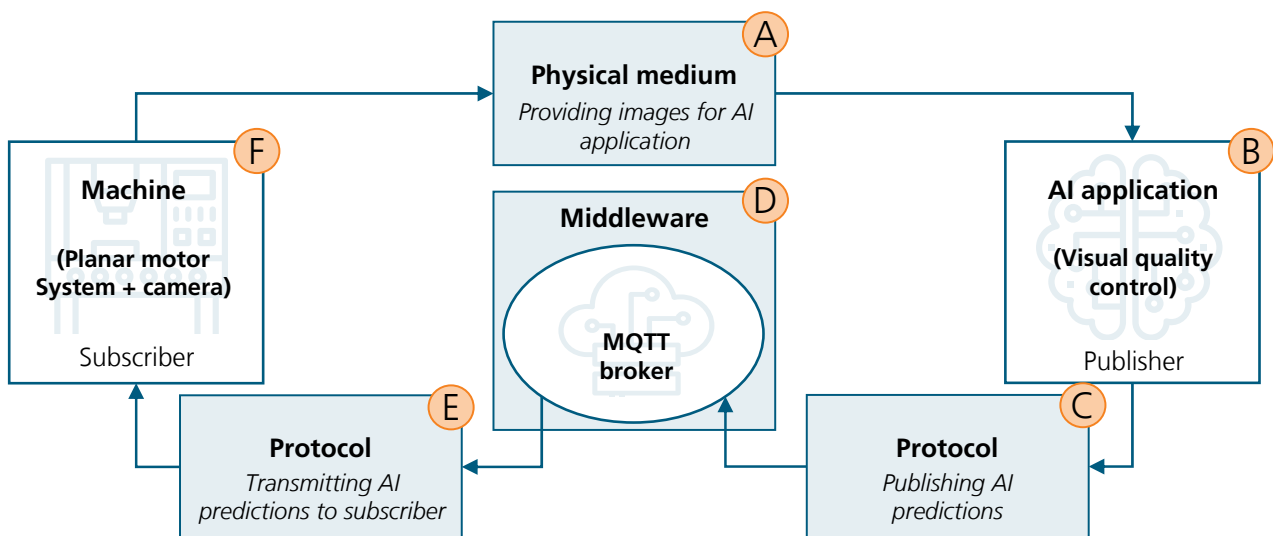


Figure 18: Demonstrator—Information and communication flow.

as crucial. A reliable transmission for example is needed, because otherwise, the control algorithm cannot accurately assign products to their designated positions. A high-speed communication is therefore not required as it is not a time-critical process but can also react after a few seconds and the machine can wait for the control algorithm's commands.

As the third step the guideline proposed MQTT as machine protocol which then was chosen for this use case.

Moreover, an MQTT broker was selected based on the selected protocol and the simplicity of connection.

An additional database is not needed because the storing of the detected data and information is not necessary for the use case. Finally, it should be mentioned that the machine is digitally connected to the internal internet, but no IIoT platform is required due to the small number of connected devices and their low complexity.

Finally in step 4 the implementation of the connection of the AI Application to the middleware and to the machine was developed.

## Conclusion and outlook

In this study, a comprehensive overview of available technologies which are relevant for the AI to shopfloor integration. Based on that key criteria were defined to support the selection process of technologies.

Moreover, a guideline for supporting the process of a seamless integration of AI applications to the shop floor was developed. The guideline includes an extensive description on important technologies which are involved in this process.

By using the results of the study companies can shorten their worktime in research and identifying the right technology. Furthermore, through the developed questionnaire companies can create more suitable solutions, based on the answered questionnaire which is aligned with the starting situation of each individual company. Additional to that the guideline also brings all relevant stakeholder together, which promotes communication and implementation of the AI application on the shopfloor.

As outlined above, the developed guideline comprised four steps, starting with the selection of a use case into which an AI application is to be seamlessly integrated and the completion of a question-and-answer catalog. These two steps helped to identify important criteria for the respective use case and laid the foundation for finding the optimal AI integration strategy tailored to this use case in the following steps. Subsequently, suitable IIoT technologies could be selected by analyzing the

provided evaluation tables and selecting the technologies that met the relevant criteria identified above. In this study, the focus was placed on machine protocols, databases and IIoT platforms as IIoT technologies. In addition, a use case was implemented in which an AI application was integrated into a production process and this guide was used to find the most suitable machine protocol.

As an outlook, the question-and-answer catalog could be expanded to include insights into which AI applications are best suited for certain machines and production scenarios. This expansion would empower users to identify the most appropriate AI tools to further refine and optimize their production processes.

# AI everywhere – Generative AI for production and business operations

---

**Liz Leutner**

Research Fellow

Production Quality

Fraunhofer Institute for Production Technology IPT

**Prof. Dr.-Ing. Robert H. Schmitt**

Member of the board of director of Fraunhofer IPT and

Holder of the chair of Production Metrology and Quality

Management of the WZL | RWTH Aachen University



# Introduction

## Motivation and objective

The rapid development and spread of Generative Artificial Intelligence (AI) in recent years has the potential to profoundly transform various areas of business and manufacturing [55]. From the automation of creative processes to the optimization of industrial production workflows - the potential applications are diverse and promising. Nevertheless, due to the novelty of the technology and the diversity of AI-generated modalities, the potential for the industry is often difficult to grasp. In order to realize the potential of Generative AI by increasing the effectiveness and efficiency of processes through the automated generation of content, this study aims to provide companies with an introduction to the world of Generative AI and its use cases in industry.

The complete results of this study were made available to ICNAP members in an interactive web application (Figure 19). It is integrated into the ICNAP Explorer [56], an interactive website to explore the projects of ICNAP. In addition to the use cases of Generative AI in an industrial context, structured

by application area, the website also provides information on the various modalities of Generative AI, available tools and the technical background. Furthermore, detailed information is provided on specific use cases that ICNAP members consider to be particularly relevant. The web application is aimed both at users who are AI beginners and want to gain an impression of this technology and its potential, as well as at technology experts who are interested in implementing specific Generative AI use cases.

This report starts with a definition and brief technical introduction to Generative AI. It lists the modalities and application areas and describes corresponding examples of use cases in the various areas. A framework for the development of Generative AI applications will then be presented before the report draws a conclusion.

## Definition of Generative Artificial Intelligence

Generative Artificial Intelligence (GenAI) describes a class of computational techniques that are able to generate seemingly new and meaningful content such as text, images or audio from training data [57]. This technology is currently revolutionizing the way we work and communicate, with examples such as DALL-E 2 [58], GPT-4 [59] and the Siemens Industrial Copilot [60].

The main models of Generative AI include different architectures like Generative Adversarial Networks (GANs) [61], Transformer models [62] and Variational Autoencoders (VAEs) [63].

The models are designed for different modalities and tasks. The training of the models is elementary and is carried out using data, through which the model learns how to generate corresponding new data. For example, the GPT (Generative Pre-trained Transformer) models are used to generate text [64]. During training, huge amounts of text data are used, whereby the model learns structures and contexts of language. To achieve this, these models require an extremely high number of parameters. For example, GPT-3 from OpenAI [65] has 175 billion parameters and was trained with a filtered Common Crawl dataset [66], a version of the WebText dataset (expanded) [67], two books corpora and Wikipedia [65].

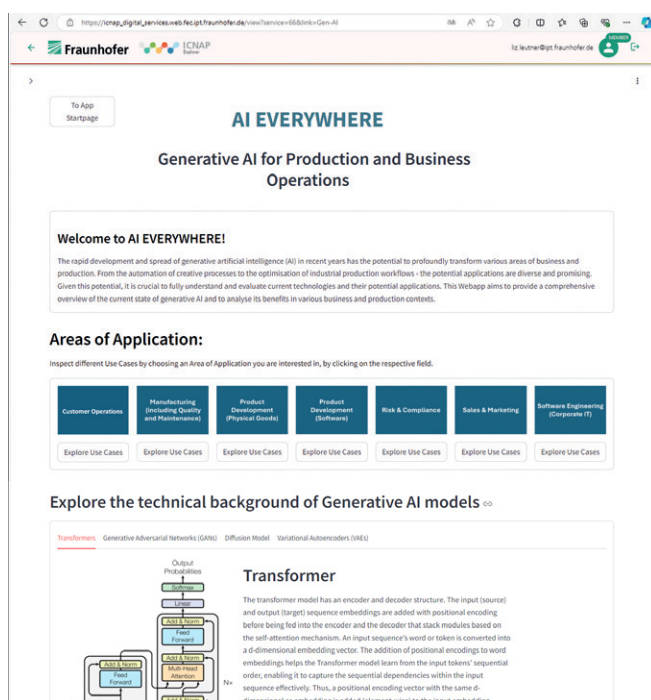


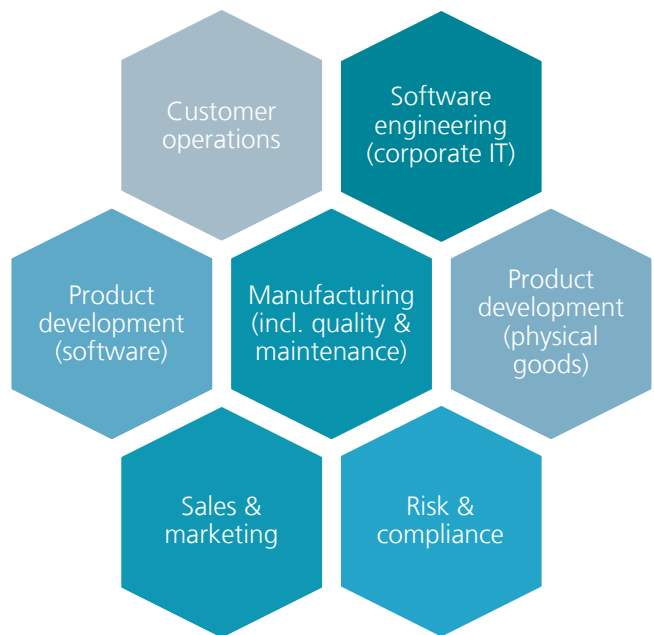
Figure 19: Excerpt from the web application for the interactive provision of information on use cases of Generative AI in an industrial context, including modalities and technical backgrounds.

## Modalities and areas of application of GenAI

The modalities of GenAI are diverse and include the generation of text, images, videos, music and even complex code [57]. However, we have additionally emphasized and differentiated modalities such as processes and CAD/3D models, as these are particularly relevant for business and production applications.

After an initial, broad research on use cases of Generative AI, we examined meta-analyses that list application areas of GenAI for the industry and evaluate them in terms of their relevance and potential [68]-[70]. Based on the insights gained, we decided on the following areas of application and assessed their relevance for the ICNAP members in order to be able to set priorities for the subsequent identification of corresponding use cases (Figure 20):

In the following section, exemplary use cases are presented for each application area in order to provide an impression of the diversity and potential of the use of Generative AI in the context of production and business operations.



*Figure 20: Overview of the application areas of GenAI for production and business operations for which use cases were identified as part of this study.*

# Selection of use cases

GenAI has shown high potential for transforming customer interactions in various business areas. A selection of exemplary use cases of Generative AI from the previously introduced application areas is presented below. First, use cases of GenAI for customer operations are considered.

## **Chatbots for personalized customer support in real-time**

One important application is the use of GenAI-enabled virtual agents like chatbots [71], [72]. These agents are designed to improve customer interaction by providing personalized support in real time. By using advanced natural language processing (NLP) and speech-to-text technologies, these virtual agents can mimic human-like qualities such as empathy and personalized communication, which are crucial for building trust and relationships with customers. This technology not only speeds up the response to customer queries, but also enables human agents to focus on more complex and differentiated issues, optimizing the allocation of resources within customer support.

## **Generation of research-based reports of customer data**

Another important application of GenAI in customer operations is the creation of research-based reports on customer data. During the onboarding process, GenAI can be used to create comprehensive reports that provide valuable insights for decision-making [73], [74]. By automating the analysis and synthesis of large volumes of customer data, this application reduces the time spent on manual research, and improves the accuracy and relevance of the information provided to employees. This leads to more informed decisions and better management of customer relationships.

## **Voice assistants for a high-quality customer experience**

In addition, GenAI is integrated into customer support through AI-powered voice assistants [72]. These assistants are able to process customer requests quickly and in accordance with company guidelines, thus maintaining or even increasing customer satisfaction. GenAI's ability to quickly process and respond to customer inquiries not only improves the efficiency of customer support, but also ensures a consistent and high-quality customer experience.

## **Software engineering (corporate IT)**

GenAI becomes increasingly important in software engineering, especially in corporate IT. It can be used to improve various stages of software development, making processes faster, more accurate, and more innovative. Use cases of this application area are described below.

## **Data management (analysis, cleaning, and labeling of large amounts of data)**

One of the primary areas where GenAI is making a difference is in data management [70], [74]. By automating the analysis, cleaning, and labeling of large amounts of data (such as user feedback and system logs) GenAI helps to process data more efficiently and accurately. This capability transforms raw data into valuable insights that supports the decision-making process and leads to more reliable software development.

## **Support in the design of IT architectures**

GenAI is also transforming the way IT architecture is designed [75], [76]. Typically, creating IT systems is a complex process that involves exploring different configurations to meet requirements like performance and security. GenAI speeds up this process by allowing engineers to quickly generate and test multiple design options. This reduces the time needed to develop systems and improves overall design quality, enabling companies to respond more quickly to changes in technology and business needs.

## **Generation of test cases and data to ensure the robustness of IT systems**

Additionally, GenAI is improving quality assurance in software development [74]. It automates the creation of test cases and data, making it easier to perform thorough testing, especially in stress scenarios where systems need to be tested under heavy load. This ensures that IT systems are more reliable and less prone to failures.

## **Product development (software)**

GenAI also plays a crucial role in enhancing the software product development process. Next, some typical use cases in this area of application are presented:

### **Automating routine coding tasks**

GenAI enhances the efficiency and consistency of software development [74]. By assisting developers in creating and maintaining multiple applications and platforms, GenAI streamlines the development process. It automates routine coding tasks, provides useful suggestions by generating code snippets, and serves as a resource for finding information quickly. This allows developers to concentrate on more complex and creative aspects of product development which leads to faster development cycles and more consistent quality across different software products.

### **Generation of test cases and data to ensure the functionality of software**

In the area of quality assurance, GenAI is transforming traditional testing processes [74]. It automates the generation of test cases and test data, making functional and performance testing more efficient and comprehensive. This automates the time-consuming manual generation of tests to check extensive code and can increase test coverage and quality. The use of GenAI in quality assurance helps in identifying potential issues early, reducing the risk of post-launch failures.

### **Content creation**

Finally, GenAI is revolutionizing content creation within software development [70], [74]. By integrating GenAI tools into content management and creation processes, it minimizes the need for manual editing and optimizes the management of large volumes of content. This is especially useful for tasks like editing videos and images, where efficiency and accuracy are critical. GenAI enables content creators to meet tight deadlines with high-quality outputs, ultimately enhancing the overall product development lifecycle.

## **Product development (physical goods)**

In the previous section the development of software products was considered. Some use cases of Generative AI for the development of physical goods are listed here:

### **Supporting virtual simulation processes**

Virtual simulations are another area where GenAI is having a transformative impact [70]. By integrating generative deep learning design techniques, GenAI significantly improves the efficiency, accuracy and innovation of simulation processes. These advanced simulations enable companies to test and refine product designs in a virtual environment, reducing the need for costly physical prototypes and accelerating the development cycle.

### **Proposing designs and materials for products**

In materials science, GenAI enables designers to explore a wider design space and optimize material properties more effectively. The process of discovering and developing new materials is inherently complex and time-consuming, but GenAI accelerates this process by identifying the most promising methods for optimizing materials and reducing the number of experiments required. This results in faster development of innovative materials with optimised properties, which ultimately improves product performance.

### **Efficient introduction of new products by automating the generation of documents**

In addition, GenAI streamlines new product inventory management by automatically creating descriptions based on existing

inventory data or user-provided information [76]. This automation can be seamlessly integrated with enterprise resource planning systems such as SAP, Oracle or Microsoft Dynamics, ensuring that product metadata is accurately and efficiently managed. This feature not only improves inventory management, but also increases the overall efficiency of product lifecycle management.

## **Manufacturing (including quality and maintenance)**

GenAI is fundamentally changing the manufacturing industry, particularly in the areas of quality, maintenance and operational efficiency. Its applications range from improving decision-making and streamlining troubleshooting processes to optimizing production and inventory management. Examples of the use of GenAI in the field of manufacturing are described here. This application area was considered particularly relevant by the ICNAP members.

### **Virtual field assistance / customized chatbots to provide real-time support**

One important use case is the integration of GenAI-enabled virtual field assistants into technical workflows [72], [74]. These virtual assistants increase operational efficiency by providing real-time support and improving decision-making processes on the factory floor. This application is particularly valuable in complex industrial environments where fast and accurate decision-making is crucial. GenAI is also being used to streamline information gathering in production environments, particularly in companies that have grown through mergers and acquisitions [70], [74]. The resulting fragmentation of systems and processes can make it difficult to quickly access the information needed. AI-powered bots solve this challenge by enabling faster and more accurate information retrieval, thus increasing employee productivity and reducing downtime.

### **Support in system diagnostics**

AI plays a crucial role in system diagnostics and maintenance by analysing system logs, user feedback and performance data [70], [74]. This analysis helps engineers to diagnose problems, suggest solutions and predict areas that need improvement, ultimately increasing the efficiency and effectiveness of maintenance work.

### **Improvement of asset maintenance planning**

Generative AI also improves asset maintenance planning [77]. By integrating AI into maintenance strategies, companies can increase asset availability, reduce costs and improve operational efficiency. This application is particularly beneficial in industries such as mining and oil and gas, where effective maintenance is critical to avoiding costly downtime and repairs.

## Sales and marketing

Generative AI has a high impact on sales and marketing, as it improves the way companies gather market knowledge, create content and plan promotions. Next, some use cases of GenAI in the application area sales and marketing are described:

### Identifying market trends by analyzing data

A key area of application for GenAI is analyzing large volumes of unstructured data, such as social media posts, news articles, research reports and customer feedback [73]. By processing this information, GenAI helps sales and marketing teams to better understand market trends and customer needs. This enables more targeted and effective communication with customers and helps companies to reach the right audience with the right message.

GenAI is also changing the way companies plan and execute trade promotions, particularly in the consumer goods sector [74]. It helps companies analyze data quickly, predict outcomes and adjust their strategies, making the promotion process more efficient and increasing the chances of success when negotiating with retailers.

### Content creation: generate marketing materials

When it comes to content creation, GenAI helps to produce consistent and personalized marketing materials, whether they are product descriptions, images, videos or audio [74]. This helps companies maintain a consistent brand message across different platforms and ensures that content is optimized for specific purposes, such as improving search engine placement or creating effective email campaigns. In addition, GenAI helps companies create marketing materials that comply with regional regulations and cultural norms [74]. This is particularly useful for companies operating in multiple countries to ensure that their marketing efforts are both effective and compliant.

### Support efficient marketing management across large and diverse product portfolios

Finally, GenAI supports the efficient management of marketing content across large and diverse product portfolios [78]. It enables companies to quickly create and update content in multiple languages to ensure consistency and a unified brand experience for customers around the world.

## Risk and compliance

Generative AI plays an important role in improving risk management and compliance in various industries, particularly in areas such as intellectual property protection, workplace safety and internal control.

Finally, examples of GenAI use cases in the area of risk and compliance are described.

### Automating the analysis of patents

In the manufacturing industry, protecting intellectual property (IP) is crucial, but often challenging due to complex patent portfolios and evolving legal frameworks. GenAI helps by automating the analysis of patents, simplifying legal processes and strengthening IP protection strategies [79]. This allows companies to navigate the complex legal landscape more efficiently and maintain solid protection for their innovations.

### Ensuring compliance with regulations through monitoring

Safety in the workplace is another area where AI is making a significant contribution [74]. AI systems can monitor and enforce safety protocols in real time, ensuring compliance with regulations such as social distancing and the use of personal protective equipment (PPE). By analyzing historical safety data, GenAI can identify potential risks and enable proactive measures to prevent accidents. This continuous monitoring helps companies maintain a safe working environment and avoid regulatory fines.

### Identifying risks through processing large amounts of data

GenAI also strengthens internal controls and corporate governance [74]. By analyzing large data sets in real time, GenAI can identify anomalies and potential risks that could indicate fraud or non-compliance. This improves transparency and accountability within organizations and supports more effective decision-making. AI-driven systems can also adapt to emerging risks, ensuring that governance frameworks evolve in line with changing business and regulatory environments.

## Deep dives

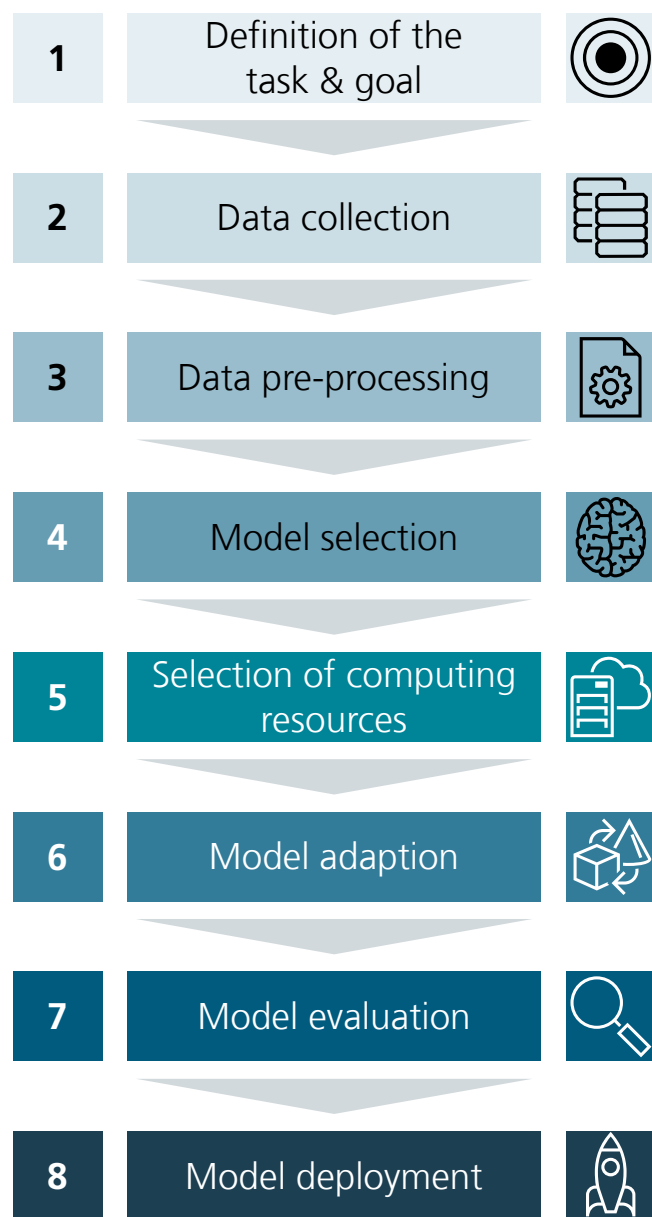
To explore specific use cases in greater detail, deep dives were conducted into three distinct areas, focusing on the foundational aspects of each application, the technical solutions involved, and their practical implementation. The objective was to gain a comprehensive understanding of potential solutions and to provide a clear overview of the current state of development in these areas. The selected deep dives are for the use cases Quality Control in Manufacturing, Process Mining with LLMs and Industrial chatbots and are accessible to ICNAP members. The same applies to the entire overview of all identified use cases.

To support the realization of these use cases, the next section introduces the most important steps in the development of corresponding GenAI applications.

# Framework for implementing GenAI applications

In this chapter, a framework that summarizes the implementation of GenAI for various use cases is presented (Figure 21). It outlines a systematic approach to the development, deployment and maintenance of GenAI models that ensures that they are both effective and consistent with best practices in data handling and model training [80]. While there are specific

approaches for specific models and use cases of GenerativeAI, the following framework is a general guideline to support first implementations [81].



## Definition of the task & goal

The process begins with problem definition and goal setting, where the specific challenge or need is identified. This includes determining the appropriate output modality (e.g. text or image) and considering key requirements such as language, resolution and style. Understanding the capabilities and limitations of the model is critical at this stage, as it helps to define the quantitative evaluation metrics that will guide the model's performance assessment [80], [81].

## Data collection

Data collection then takes place to gather the necessary information from reliable sources such as databases, web scraping or APIs [82]. Ensuring data quality and diversity is crucial, as is compliance with legal and ethical standards, especially when dealing with copyrighted or sensitive information. Data protection laws must be observed, and data should be stored securely, with encryption and appropriate anonymization [80], [83].

## Data pre-processing

The collected data is then pre-processed, where it is cleaned to remove inconsistencies, normalized to ensure consistent scales, and augmented to increase the robustness of the dataset. Accurate labelling is essential for supervised learning tasks, and the data is split into training, validation and test sets to support effective model development [80].

## Model selection

At the same time, a foundation model is selected based on the specific task, the compatibility of the dataset and the computational requirements. Popular models such as GPT-4, LLaMA or Google Gemini are considered, paying attention to their transfer learning capabilities and community support to ensure that they meet the project's requirements [81], [84].

## Selection of computing resources

In parallel, the selection of hardware or cloud service is crucial to meet the model's requirements in terms of computing power, memory and storage. The choice of hardware – whether CPUs, GPUs or TPUs – depends on the complexity of the model, while cloud services from providers such as AWS, GCP or Azure are evaluated for cost efficiency, scalability and compatibility with machine learning tools [80].

Figure 21: Framework for the development of Generative AI applications. It can serve as a general guideline for initial developments.



### Model adaption

Next, the model undergoes training, fine-tuning and retrieval augmented generation (RAG). In this phase, the weights and parameters of the model are adapted to the respective tasks, with regularization techniques being used to prevent overfitting [85]. RAG is implemented to improve the generative results by retrieving relevant information and integrating it into the model's responses to provide richer, contextually informed responses [80].

### Model evaluation

This is followed by the evaluation phase, where metrics such as FID, BLEU and ROUGE are used to assess the performance of the model [85]. Validation and test sets are crucial here for fine-tuning the hyperparameters and ensuring good generalisation of the model. In addition, qualitative analysis helps to identify and correct biases or errors, and feedback loops are set up for ongoing refinement and monitoring of data deviation [83].

### Model deployment

Finally, the model is deployed to ensure that it is operational in a real-world environment. This includes setting up the necessary hardware infrastructure, using containerisation tools such as Docker and orchestration systems such as Kubernetes for consistent deployment [86]. The model is integrated with APIs for application access, and ongoing monitoring ensures that key performance indicators are met, ethical considerations are taken into account and security measures are in place to protect the deployment infrastructure.

This framework provides a comprehensive overview of the key steps in implementing GenAI applications and offers a structured approach to ensure that these models are used in a way that is both effective and efficient in a range of use cases [80]. The framework is not exclusively intended to support the development of industrial GenAI applications but can also be used in this area to optimize the development of corresponding applications for the optimization of processes from production and business operations. It is particularly suitable for supporting the development of proof-of-concepts or prototypes for a specific use case that is relevant for an enterprise in order to gain an impression of the suitability of GenAI for fulfilling the specific requirements of the company and the task.

## Conclusion

This study provides an introduction to the world of Generative AI in the field of production and business operations. Technical backgrounds were described and structured use cases were introduced in various areas. A framework for the implementation of Generative AI applications was also presented. This helps companies to gain an impression of the potential of this new technology and to identify and implement relevant use cases according to their own requirements.

Nevertheless, the challenges of using GenAI in industrial contexts must also be considered. These include the lack of explainability and transparency of model outputs, the production of false information (hallucination) by language models,

the high computing capacity often required and the lack of availability of high-quality data and data sources in an industrial context, as well as unresolved issues relating to data security [41], [87], [88]. Future research must address these challenges in order to support the industrial use of GenAI.

# Towards a Dark Factory – Leveraging multidimensional twins in a manufacturing metaverse

---

**David Wichter**

Research Fellow  
High Performance Cutting  
Fraunhofer Institute for Production Technology IPT

**Gisele Sà Lima Cruz**

Student Assistant  
High Performance Cutting  
Fraunhofer Institute for Production Technology IPT

**Prof. Dr.-Ing. Thomas Bergs**

Member of the board of director of Fraunhofer IPT and  
holder of the chair of Manufacturing Technology  
Management of the MTI of RWTH Aachen University.



# Introduction

The history of industry can be understood as the advancement of the way that humans produce goods and services. Throughout this progress several industrial revolutions again and again harnessed new energy sources, reorganized labor, and ultimately increased productivity. Today new market pressures such as the trend for mass customization and climate change force businesses to innovate further [89]. The fourth industrial revolution is a means to overcome these challenges by thoroughly connecting production resources through digital technologies. This revolution is still well underway and focus of many research activities worldwide [90]. (Figure 22)

In the wake of the current fourth industrial revolution the concept of the Smart Factory was introduced, i.e. a factory that is highly automated and smartly connected [91]. By taking this idea one step further, one arrives at the concept of a so-called Dark Factory, also called a lights-out factory [92], [93]. In this kind of factory, the degree of automation is so high, that human oversight is no longer necessary, and the lights can be

turned off, hence the name. Such a fully automated production facility, provided that it can retain or even increase flexibility, poses a substantial advantage in the modern competitive production environment. Not only is it able through smart communication among its resources to handle small lot sizes and customized, heterogenous production orders [94], [95], but it also has a higher production rate and is highly efficient. Automation to such an extent has even been described as “the holy grail of manufacturing” [92]. It is therefore an attractive goal to work towards this vision for producing businesses.

However, the idea of the Dark Factory is yet to be widely implemented and the path towards its implementation seems unclear. Businesses that want to work towards achieving the vision can find themselves quickly lost in a wide variety of different products and technologies. This study aims to alleviate this. It offers guidance by providing information and tools to answer the following questions for prospective Dark Factories:

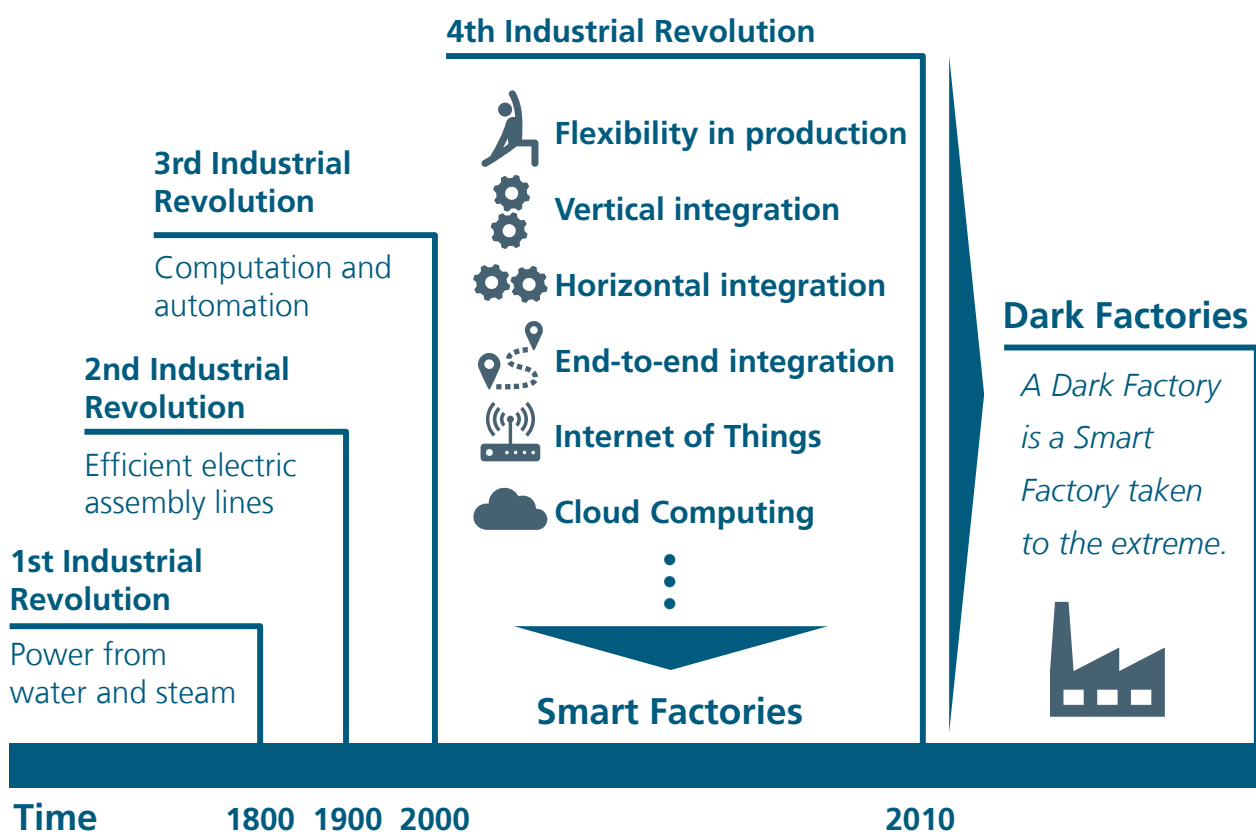


Figure 22: Smart and Dark Factories in the context of past and current industrial revolutions.

1. How close is a given factory to the goal of the Dark Factory?
2. Which technologies are or will be relevant for implementing the vision of the Dark Factory?

Taken together these questions can influence the direction of development efforts towards the Dark Factory. Section The next section explains the methodology and structure of the

study to answer these questions, and describes which deliverables were elaborated. The deliverables are then explained in more detail and highlight how businesses can use them. The report concludes with a short summary of how the study could support companies on their way towards the vision of the Dark Factory.

## Study deliverables and methodology

The research for this study was conducted in three phases with two main deliverables as results. The employed procedure is visualized in Figure 23.

The three phases of the study were as follows:

**Phase I:** First a literature review was conducted, in which sources concerning the Dark Factory were gathered in a

catalogue and evaluated. Then they were used to extract the dimensions of the maturity model, i.e. the aspects in which factories progress towards a Dark Factory.

**Phase II:** In the second phase the maturity model was finalized by adding stages and capabilities to the dimensions. Furthermore, concrete technologies were researched from literature as well as from exchanges with the community that for each dimension help achieve these capabilities. They were additionally mapped on to the ICNAP topic fields.

**Phase III:** In the last phase time horizons were defined that represent categories for the future availability of a technology. Each technology was then assigned to these time horizons in order to draw the roadmap for the Dark Factory.

Over the course of the study two main deliverables were achieved:

**A maturity model:** This model includes the dimensions, stages, and capabilities of a Dark Factory and captures the difference in development between traditional factories and smart, highly autonomous Dark Factories. It is a tool to assess for a given factory how close it is to the goal of the true Dark Factory in terms of capabilities. It can be used to deduce in which dimensions more work needs to be done .

**A roadmap:** In order to focus efforts in the right direction and not miss important research trends, the roadmap can be used. It organizes technologies that are important in the Dark Factory context into time horizons and thus gives them a chronological order. By using this tool future technologies and competencies can be acquired at the right time as they are needed in order to reach the capabilities that are specified within the maturity model. To that end the roadmap is also organized along the same dimensions as the latter.

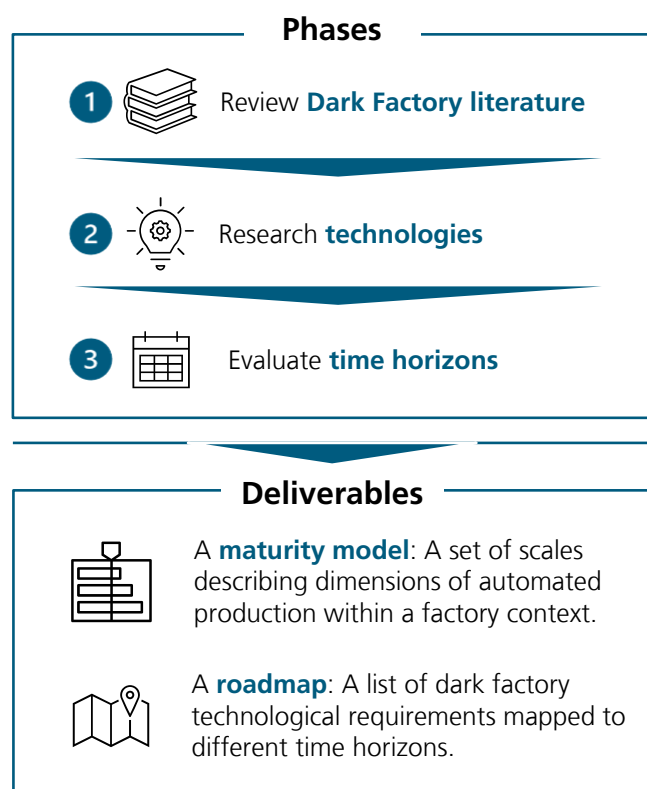


Figure 23: Phases and deliverables of the Dark Factory study.

# Maturity model

The maturity model consists of six dimensions that represent independent aspects of a factory, in each of which progress towards a Dark Factory can be made. Each dimension is partitioned into stages, which are to be understood as a level a factory can reach for that given dimension. In order to reach such a stage a factory needs to support the capabilities associated

with the given stage as well as the capabilities of the stages below it. The model is shown in Figure 24.

In the following, the model's dimensions, stages, and capabilities are described.

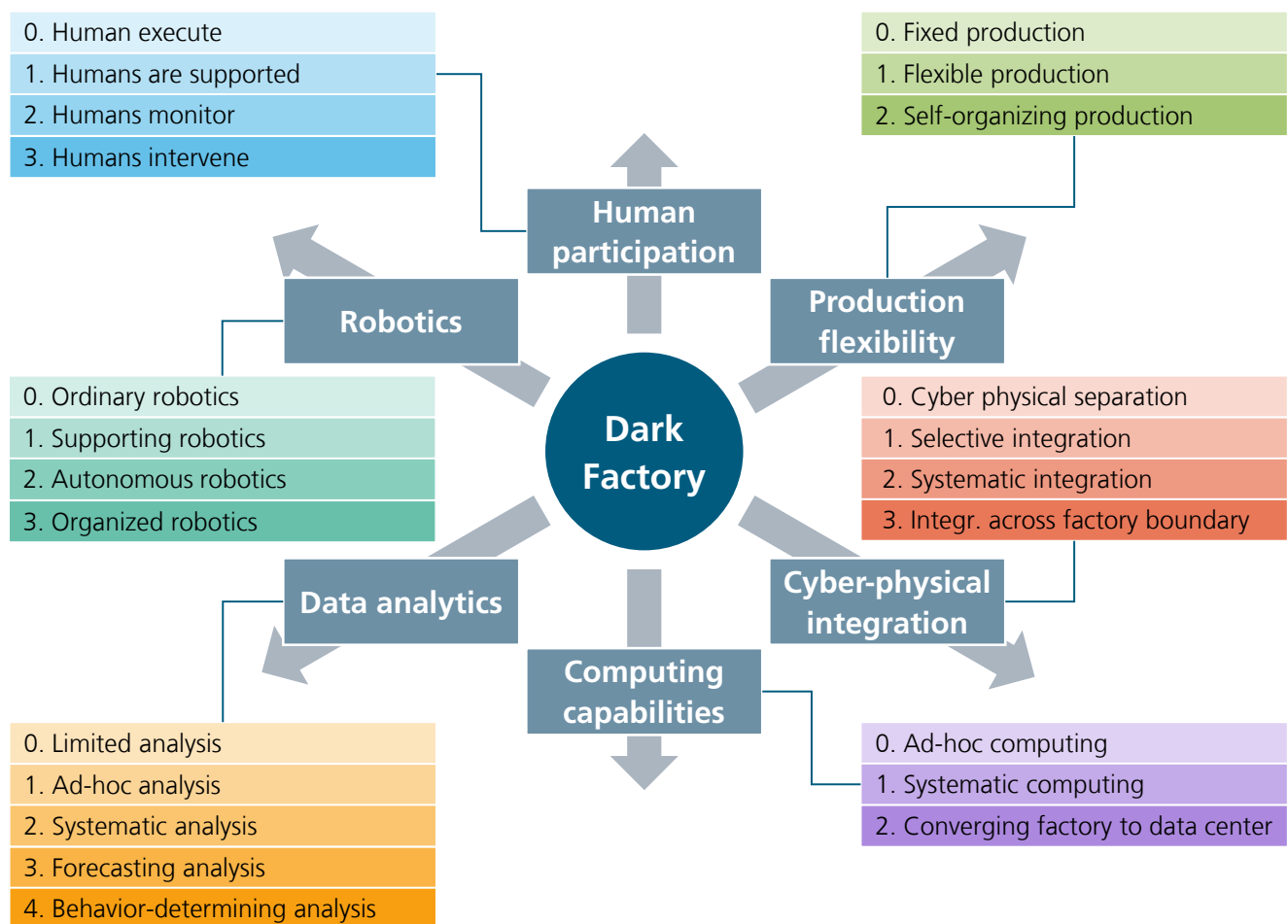


Figure 24: The dimensions and stages of the Dark Factory Maturity Model .

**Human participation:** This dimension describes the extent to which humans take part in the production process, i.e. whether they act executively or in a managing capacity. While the motivation behind the Dark Factory envisions a future without any human involvement, it remains unclear whether this will come to pass. However, as the push for further automation progresses the role of human labor will change. The stages of this dimension are as follows:

1. **Stage 0 – Humans execute:** The majority of the work is done through human labor. There are no extraordinary capabilities on this stage.
2. **Stage 1 – Humans are supported:** While a significant portion of the workload is still handled by humans, the factory supports them through functions such as remote access, predictive maintenance planning and proposals for self-optimization.
3. **Stage 2 – Humans monitor:** The majority of the workload is handled by automated systems of the factory, with humans playing a supporting role. The capabilities here include remote control, automated procurement, automated work setup and automated execution of production.
4. **Stage 3 – Humans intervene:** The factory runs autonomously, and humans only have to intervene in case of serious faults and emergencies. At this stage the factory supports functions such as automated maintenance and fully automatic self-optimization.

**Production flexibility:** This dimension expresses how well the factory handles process changes, e.g. due to small lot sizes. With the trend for mass customization being a major driver behind the Industry 4.0 movement [94], it remains important that factories will be able to meet this trend. Therefore, it is not enough to increase the degree of automation in a mass production scenario, but also to apply the same level of automation to highly customized products and parts. A factory can reach the following stages within this dimension:

1. **Stage 0 – Fixed production:** Within the factory it is difficult and expensive to change processes and produce new products. There are no extraordinary capabilities associated with this stage.
2. **Stage 1 – Flexible production:** Processes that are executed within the factory can be changed quickly and with relatively low efforts. Functions at this level include the ability to manufacture multiple different products on the same production line and active process planning support from intelligent systems within the factory.
3. **Stage 2 – Self-organizing production:** On this stage the systems within the factory plan and execute production processes autonomously based on given requirements. The lot size plays a negligible role and most if not all aspects of planning, such as processes, logistics and shopfloor management are organized by the factory itself.

**Cyber-physical integration:** This dimension describes the level of integration between the realms of the physical and the virtual, e.g. in the context of digital twins or Cyber-Physical Systems. By leveraging the strengths of virtual and software components, a factory can benefit from smarter and more connected systems, which in turn can run the factory more autonomously. Thus, the progress along this dimension constitutes progress towards a Dark Factory. In detail the following stages were identified along this dimension:

1. **Stage 0 – Cyber-physical separation:** The physical and the virtual realm are not connected within the factory, therefore no capabilities are available in this regard.
2. **Stage 1 – Selective integration:** Systems that integrate the physical and the virtual are in use on a per-need basis, such as product digital twins or Cyber-Physical Systems in the form of machine tools.
3. **Stage 2 – Systematic integration:** Not only are digital twins and Cyber-Physical Systems commonplace for products and machines, but they are also integrated with each other and communicate. Important aspects of the factory can be controlled thus by interfacing with the virtual world.
4. **Stage 3 – Integration across factory boundaries:** The factory is integrated beyond its own inner system in a virtual ecosystem, for example like the metaverse.

**Computing capabilities:** This dimension assesses the capabilities of the factory in terms of computation, e.g. on-premises, on the edge, or in the cloud. As the need for computational resources increases with the use of more sophisticated data processing techniques and the handling of higher data volumes, the data center and the factory converge more and more. The dimension of computing capabilities represents this development and consists of the following stages:

1. **Stage 0 – Ad-hoc computing:** There are no extraordinary computational resources in use within the factory.
2. **Stage 1 – Systematic computing:** The factory makes extended use of computing facilities, with a mixture of on-premises, edge, and cloud computing capabilities. Within its capabilities is also secure data transmission, storage and processing.
3. **Stage 2 – Factory to data center convergence:** The factory supports the functions of a high-end data center. The capabilities include hyperscaling of resources, big data and real-time support.

**Data analytics:** This dimension gauges the degree of how much of decision-making is supported by data analytic techniques, such as AI and machine learning. As data is not an end in itself but rather a tool to make informed decisions, methods and tools are needed to gain knowledge from the data. The discipline of data analytics can be used to accomplish this and is therefore an integral part of a Dark Factory. The stages are:

1. **Stage 0 – Limited analysis:** Data analytics are not used in a meaningful way.
2. **Stage 1 – Ad-hoc analysis:** Simple analytics methods are used on a per-need basis. They mostly concern the area of descriptive data analytics.
3. **Stage 2 – Systematic analysis:** Data analytics on this stage is an automated process that leads to systematic insights about the production process. The capabilities are of the realm of diagnostic data analytics.
4. **Stage 3 – Forecasting analysis:** Within the factory predictive data analytics is used to make estimations about the future which inform the behavior and the decisions during production.
5. **Stage 4 – Behavior-determining analysis:** Prescriptive data analytics is used to govern behavior autonomously within the factory.

**Robotics:** This dimension signifies the extent as to which human labor is replaced by cooperative or autonomous robots. As human involvement decreases on the path towards more automation, the physical labor needs to be taken up by machines and robots. A factory therefore needs to progress along this dimension on its way to the Dark Factory. In doing so it reaches the following stages:

1. **Stage 0 – Standard robotics:** There are no special abilities concerning robotics within the factory that go beyond basic functions.
2. **Stage 1 – Supporting robotics:** Robots are actively used to support human workers. Examples include cobots and autonomous guided vehicles (AGVs).
3. **Stage 2 – Autonomous robotics:** At this stage robots within the factory act as autonomous agents that fulfill general tasks assigned by human supervisors.
4. **Stage 3 – Organized robotics:** Ultimately, the robots within a Dark Factory are self-organized, with little or no human intervention.

# Roadmap

The roadmap identifies technologies, trends and research topics that are relevant for the foreseeable future as factories evolve towards more automation and autonomous behavior. It can indicate where research efforts should be placed and what the expected technological advancement in a given Dark Factory might look like. Naturally the more the timeframe for predictions stretches into the future, the larger the cone of uncertainty grows. Making confident assumptions thus quickly becomes difficult. To meet this problem three different time horizons have been defined to indicate how quickly a given technology is expected to be widely available. These time horizons are as follows:

**Market ready:** As the name implies, technologies that fall within this horizon are already widely available and in use by many market players. The technology readiness level [96] for these technologies falls within level 8 and 9. Thus there is at least one product or service on the market that implements the given technology. These products help factories progress along the dimensions to reach the stages and capabilities described in the maturity model. Some examples are listed in Table 3. The references for each technology are examples that highlight their active use within manufacturing and elsewhere.

**Short term:** Technologies that are already proven to be promising in the context of Dark Factories, either through working prototypes or extensive research, fall within this time horizon. As a rule, a technology can be classified in this area if commercial solutions are foreseeable, for example because a startup is working on a corresponding product, or if application-oriented research has proven the feasibility of the technology. Therefore, these technologies generally have a readiness level between 4 and 7 and can be expected to be available in the upcoming years. Table 4 lists some examples. Here, the references illustrate technologies that are actively being researched, with market readiness in short term reach.

**Long term:** Finally, there are technology trends and topics that are expected to play an important role in the development of the vision of the Dark Factory. However, they are not yet market ready or still in early stages in the context of production. Their technology readiness levels are between 1 and 3 which means they are either in the proof-of-concept or early research phase. Some of these trends are listed in Table 5. The references show that these technologies are currently conceived but not soon to be applied in real production environments.

Dimension	Technologies and trends
Human participation	Secure data governance [97], remote commissioning [98] and maintenance [99], pay-per-use [100] and pay-per-part [101]
Production flexibility	Wifi 5G and 6 [102], resource hyperscaling [103], predictive maintenance [104]
Cyber-physical integration	Augmented [105] and virtual [106] reality, ready-made DT tools [107], real-time tracking [108], standardized protocols [109], IoT databases [110]
Computing capabilities	Cloud [111] & edge [112] computing, hybrid cloud storage [113], containerization [114], secure cloud architectures [115], secure data at rest [116] and in transit [117], custom trained AI [118], big data processing [119], Industrial Internet of Things [120], environmental sensing [121], computational sensors [122]
Data analytics	Generative AI [123], semantic data models [124], standardized information models [125], high frequency data acquisition [126]
Robotics	Cobots [127], automated guided vehicles [128], autonomous mobile robots [129], tracking and tracing [130], M2M communication [131]

Table 3: Market ready roadmap trends and technologies.

Dimension	Technologies and trends
Human participation	Emotion recognition [132], gesture-based interfaces [133], human oversight in AI [134]
Production flexibility	Manufacturing-as-a-service [135], cloud manufacturing [136], Smart retrofitting [137]
Cyber-physical integration	DT standards [138], physical process models [139], wireless sensing [140], DT monetization [141], real-time DT [142]
Computing capabilities	Real-time data acquisition [143], AI-on-device [144], DT-on-edge [145]
Data analytics	Data monetization [146], data exchange spaces [147], AI-integrated DTs [148], AI integrated sensors [149], explainable AI [150]
Robotics	Plug-and-Play Robotics [151], cloud-Integrated robotics [152], cyber-physical systems [153]

Table 4: Short term roadmap trends and technologies.

Dimension	Technologies and trends
Human participation	Complete horizontal and vertical integration [154]
Production flexibility	Self-organized systems [155], AI-based production planning and scheduling [156]
Cyber-physical integration	Metaverse [157], ambient IoT [158]
Computing capabilities	IT/OT-convergence [159], quantum computing [160], virtual control [161]
Data analytics	Qualified (trustworthy) AI [162]
Robotics	Human-level general purpose robotics [163]

Table 5: Long term roadmap trends and technologies.

## Conclusion

Dark Factories hold many promises for future production scenarios, with higher productivity at the top of the list. Whether or not the concept can fulfil these promises depends on whether it can be widely and safely implemented. To that end a path needs to be drawn that makes the implementation feasible. With the outcomes of this study drawing this path for a given organization has become easier. The proposed maturity

model and technology roadmap can indicate the progress of a factory towards a Dark Factory, which can help companies identify room for development within their own production facilities. Companies are enabled to formulate their own vision of a Dark Factory and receive guidance during the definition of steps towards this defined vision.

# References

- [1] V. Rizos and P. Urban. "IMPLEMENTING THE EU DIGITAL BATTERY PASSPORT: Opportunities and challenges for battery circularity." Accessed: Oct. 24, 2024. [Online]. Available: [https://circulareconomy.europa.eu/platform/sites/default/files/2024-03/1qp5rxiz-CEPS-InDepthAnalysis-2024-05\\_Implementing-the-EU-digital-battery-passport.pdf](https://circulareconomy.europa.eu/platform/sites/default/files/2024-03/1qp5rxiz-CEPS-InDepthAnalysis-2024-05_Implementing-the-EU-digital-battery-passport.pdf)
- [2] European Parliament. "Circular economy: definition, importance and benefits: The circular economy: find out what it means, how it benefits you, the environment and our economy." Accessed: Oct. 24, 2024. [Online]. Available: <https://www.europarl.europa.eu/topics/en/article/20151201STO05603/circular-economy-definition-importance-and-benefits>
- [3] Volkswagen. "Recycling und Rücknahme." Accessed: Oct. 24, 2024. [Online]. Available: <https://www.volkswagen.de/de/besitzer-und-service/ueber-ihr-auto/kundeninformationen/rechtliches/recycling-und-ruecknahme.html>
- [4] European Environment Agency. "life cycle assessment." Accessed: Oct. 24, 2024. [Online]. Available: <https://www.eea.europa.eu/help/glossary/eea-glossary/life-cycle-assessment>
- [5] R. Wintergerst. "Wirtschaftsschutz 2024." Accessed: Nov. 6, 2024. [Online]. Available: <https://www.bitkom.org/sites/main/files/2024-08/240828-bitkom-charts-wirtschaftsschutz-cybercrime.pdf>
- [6] Bitkom. "Angriffe auf die deutsche Wirtschaft nehmen zu." Accessed: Sep. 3, 2024. [Online]. Available: <https://www.bitkom.org/Presse/Presseinformation/Wirtschaftsschutz-2024>
- [7] A. Fleck. "Cybercrime Expected To Skyrocket in Coming Years." Accessed: Sep. 10, 2024. [Online]. Available: <https://www.statista.com/chart/28878/expected-cost-of-cybercrime-until-2027/>
- [8] Allianz. "Allianz Risk Barometer 2024 - Rank 1: Cyber incidents." Accessed: Sep. 7, 2024. [Online]. Available: <https://commercial.allianz.com/news-and-insights/expert-risk-articles/allianz-risk-barometer-2024-cyber-incidents.html>
- [9] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, and S. Lightman, "Guide to Operational Technology (OT) Security: Initial Public Draft. NIST Special Publication. NIST SP 800-82r3 ipd," 2022, doi: 10.6028/NIST.SP.800-82r3.ipd.
- [10] Bundesamt für Sicherheit in der Informationstechnik. "Positionspapier Zero Trust 2023." Accessed: Oct. 10, 2024. [Online]. Available: [https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeLeitlinien/Zero-Trust/Zero-Trust\\_04072023.pdf?\\_\\_blob=publicationFile&v=4](https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/TechnischeLeitlinien/Zero-Trust/Zero-Trust_04072023.pdf?__blob=publicationFile&v=4)
- [11] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture: NIST Special Publication 800-207," 2020, doi: 10.6028/NIST.SP.800-207.
- [12] Bundesamt für Sicherheit in der Informationstechnik, *IT-Grundschutz-Kompendium*. Köln: Reguvis, 2023.
- [13] Bitkom. "Cybersicherheit: Deutscher Markt erstmals über 10-Milliarden-Marke." Accessed: Sep. 5, 2024. [Online]. Available: <https://www.bitkom.org/Presse/Presseinformation/Cybersicherheit-Deutscher-Markt-ueber-10-Milliarden-Marke>
- [14] Cybersecurity and Infrastructure Security Agency. Cybersecurity Division. "Zero Trust Maturity Model." Accessed: Nov. 6, 2024. [Online]. Available: [https://www.cisa.gov/sites/default/files/2023-04/zero\\_trust\\_maturity\\_model\\_v2\\_508.pdf](https://www.cisa.gov/sites/default/files/2023-04/zero_trust_maturity_model_v2_508.pdf)
- [15] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, "Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook," *IEEE Access*, vol. 8, pp. 220121–220139, 2020, doi: 10.1109/ACCESS.2020.3042874.
- [16] B. El-Jawhari, S. Halbe, M. Whyte, K. Cobbaert, and A. Odenkirchen, "An introduction to implementing AI in manufacturing," 2020. Accessed: Sep. 18, 2024. [Online]. Available: <https://www.pwc.com/gx/en/industrial-manufacturing/pdf/intro-implementing-ai-manufacturing.pdf>
- [17] F. Tao and M. Zhang, "Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing," *IEEE Access*, vol. 5, pp. 20418–20427, 2017, doi: 10.1109/ACCESS.2017.2756069.
- [18] S. Mantravadi, A. D. Jansson, and C. Møller, "User-Friendly MES Interfaces: Recommendations for an AI-Based Chatbot Assistance in Industry 4.0 Shop Floors," in *Intelligent Information and Database Systems: 12th Asian Conference, ACIIDS 2020, Phuket, Thailand, March 23–26, 2020, Proceedings, Part II* (Lecture Notes in Computer Science 12034), N. T. Nguyen, K. Jearanaitanakij, A. Selamat, B. Trawiński, and S. Chittayasothorn, Eds., Cham: Springer, 2020, pp. 189–201.



- [19] D. Obodovski. "IIoT's Industrial Internet Connectivity Framework and Why You Should Care." Accessed: Sep. 10, 2024. [Online]. Available: <https://www.iotworldtoday.com/iiot/iic-s-industrial-internet-connectivity-framework-and-why-you-should-care>
- [20] A. E. H. Gabsi, "Integrating artificial intelligence in Industry 4.0: insights, challenges, and future prospects—a literature review," *Ann. Oper. Res.*, 2024, doi: 10.1007/s10479-024-06012-6.
- [21] Á. Huertas-García, C. Martí-González, R. G. Maezo, and A. E. Rey, "A Comparative Study of Machine Learning Algorithms for Anomaly Detection in Industrial Environments: Performance and Environmental Impact," in *Trends in Sustainable Computing and Machine Intelligence: Proceedings of ICTSM 2023* (Algorithms for Intelligent Systems), S. Lanka, A. Sarasa-Cabezuelo, and A. Tugui, Eds., Singapore: Singapore, 2024, pp. 373–389.
- [22] R. Dzhusupova, J. Bosch, and H. H. Olsson, "Choosing the right path for AI integration in engineering companies: A strategic guide," *J. Syst. Softw.*, vol. 210, 2024, Art. no. 111945, doi: 10.1016/j.jss.2023.111945.
- [23] K. T. P. Nguyen, K. Medjaher, and D. T. Tran, "A review of artificial intelligence methods for engineering prognostics and health management with implementation guidelines," *Artif. Intell. Rev.*, vol. 56, no. 4, pp. 3659–3709, 2023, doi: 10.1007/s10462-022-10260-y.
- [24] N. Dipsis and K. Stathis, "A RESTful middleware for AI controlled sensors, actuators and smart devices," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, no. 7, pp. 2963–2986, 2020, doi: 10.1007/s12652-019-01439-3.
- [25] M. Elkhodr, S. Khan, and E. Gide, "A Novel Semantic IoT Middleware for Secure Data Management: Blockchain and AI-Driven Context Awareness," *Future Internet*, vol. 16, no. 1, 2024, Art. no. 22, doi: 10.3390/fi16010022.
- [26] J. Zhang, M. Ma, P. Wang, and X. Sun, "Middleware for the Internet of Things: A survey on requirements, enabling technologies, and solutions," *J. Syst. Archit.*, vol. 117, 2021, Art. no. 102098, doi: 10.1016/j.sysarc.2021.102098.
- [27] Patrick. "Datenbanken: Kompakt erklärt." Accessed: Sep. 10, 2024. [Online]. Available: <https://www.alexanderthamm.com/de/blog/datenbanken-kompakt-erklart/>
- [28] N. Fatima. "Verschiedene Arten von Datenbanken im Jahr 2024: Ein umfassender Leitfaden." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.astera.com/de/type/blog/a-quick-overview-of-different-types-of-databases/>
- [29] IONOS Redaktion. "Datenbanken: Wozu man sie braucht und welche Arten es gibt." Accessed: Sep. 10, 2024. [Online]. Available: <https://www.ionos.de/digitalguide/hosting/hosting-technik/datenbanken/>
- [30] K. Pykes. "What is A Graph Database? A Beginner's Guide." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.datacamp.com/blog/what-is-a-graph-database>
- [31] A. Hazra, M. Adhikari, T. Amgoth, and S. N. Srirama, "A Comprehensive Survey on Interoperability for IIoT: Taxonomy, Standards, and Future Directions," *ACM Comput. Surv.*, vol. 55, no. 1, 2023, Art. no. 9, doi: 10.1145/3485130.
- [32] Amazon Web Services, Inc. "AWS IoT: Entsperren Sie Ihre IoT-Daten und beschleunigen Sie das Unternehmenswachstum." Accessed: Sep. 17, 2024. [Online]. Available: <https://aws.amazon.com/de/iot/?nc=s&loc=0>
- [33] S. George. "Microsoft Azure IoT Suite – Connecting Your Things to the Cloud." Accessed: Sep. 17, 2024. [Online]. Available: <https://azure.microsoft.com/de-de/blog/microsoft-azure-iot-suite-connecting-your-things-to-the-cloud/>
- [34] Siemens. "Insights Hub." Accessed: Sep. 17, 2024. [Online]. Available: <https://plm.sw.siemens.com/de-DE/insights-hub/>
- [35] V. Saranya, M. J. Carmel Mary Belinda, and G. R. Kanagachidambaresan, "An Evolution of Innovations Protocols and Recent Technology in Industrial IoT," in *Internet of Things for Industry 4.0: Design, Challenges and Solutions* (EAI/Springer Innovations in Communication and Computing), G. R. Kanagachidambaresan, R. Anand, E. Balasubramanian, and V. Mahima, Eds., Cham: Springer, 2020, pp. 161–175.
- [36] D. L. Tran, T. Yu, and M. Riedl, "Integration of IIoT Communication Protocols in Distributed Control Applications," in *Proceedings. IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society: Online. Singapore. 19 - 21 October, 2020*, 2020, pp. 2201–2206, doi: 10.1109/IECON43393.2020.9254220.
- [37] A. Busboom, "Automated generation of OPC UA information models — A review and outlook," *J. Ind. Inf. Integration*, vol. 39, 2024, Art. no. 100602, doi: 10.1016/j.jii.2024.100602.
- [38] S. Figueroa-Lorenzo, J. Añorga, and S. Arrizabalaga, "A Survey of IIoT Protocols: A Measure of Vulnerability Risk Analysis Based on CVSS," *ACM Comput. Surv.*, vol. 53, no. 2, 2021, Art. no. 44, doi: 10.1145/3381038.
- [39] K.-J. Kwak and J.-M. Park, "A Study on Semantic-Based Autonomous Computing Technology for Highly Reliable Smart Factory in Industry 4.0," *Appl. Sci.*, vol. 11, no. 21, 2021, Art. no. 10121, doi: 10.3390/app112110121.

- [40] F. Ercan, M. Bega, and B. Kuhlenkötter, "A systematic literature review of communications standards in discrete manufacturing," in *Conference on Production Systems and Logistics: Stellenbosch Institute for Advanced Study (STIAS), Stellenbosch, South Africa. 14th – 17th November 2023 Proceedings*, D. Herberger and M. Hübner, Eds., 2023, pp. 80–89, doi: 10.15488/15260.
- [41] J. Krauß, T.-H. Hülsmann, L. Leyendecker, and R. H. Schmitt, "Application Areas, Use Cases, and Data Sets for Machine Learning and Artificial Intelligence in Production," in *Production at the Leading Edge of Technology: Proceedings of the 12th Congress of the German Academic Association for Production Technology (WGP), University of Stuttgart, October 2022* (Lecture notes in production engineering), M. Liewald, A. Verl, T. Bauernhansl, and H.-C. Möhring, Eds., Cham: Springer, 2023, pp. 504–513.
- [42] M. Nast, H. Raddatz, B. Rother, F. Golatowski, and D. Timmermann, "A Survey and Comparison of Publish/Subscribe Protocols for the Industrial Internet of Things (IIoT)," in *Proceedings of the 12th International Conference on the Internet of Things 2022: Delft, Netherlands. November 7-10, 2022*, E. Niforatos, G. Kortuem, N. Meratnia, J. Siegel, and F. Michahelles, Eds., 2022, pp. 193–200, doi: 10.1145/3567445.3571107.
- [43] B. U. Deveci, H. Bas, E. Ummak, O. Albayrak, and P. Unal, "A Thorough Analysis and Comparison of Data Communication Protocols Used in Industry 4.0: the Case of Smart-CNC," in *2022 International Conference on Future Internet of Things and Cloud. Proceedings: 22-24 August 2022. Hybrid Event. Rome Italy*, M. Younas, I. Awan, and W. Rahayu, Eds., 2022, pp. 199–206, doi: 10.1109/FiCloud57274.2022.00034.
- [44] E. Trunzer, P. Prata, S. Vieira, and B. Vogel-Heuser, "Concept and Evaluation of a Technology-independent Data Collection Architecture for Industrial Automation," in *Proceedings. IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society: Convention Center. Lisbon, Portugal. 14 - 17 October, 2019*, 2019, pp. 2830–2836, doi: 10.1109/IECON.2019.8927399.
- [45] D. Hastbacka, L. Barna, M. Karaila, Y. Liang, P. Tuominen, and S. Kuikka, "Device status information service architecture for condition monitoring using OPC UA," in *19th IEEE International Conference on Emerging Technologies and Factory Automation: September 16-19, 2014, Barcelona, Spain*, 2014, doi: 10.1109/ETFA.2014.7005141.
- [46] A. Ahmad. "Why Is It Hard to Horizontally Scale SQL Databases?: Understanding the Challenges of Horizontally Scaling SQL Databases for System Design." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.designgurus.io/blog/horizontally-scale-sql-databases>
- [47] AltexSoft. "Comparing Database Management Systems: MySQL, PostgreSQL, MSSQL Server, MongoDB, Elasticsearch, and others." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.altexsoft.com/blog/comparing-database-management-systems-mysql-postgresql-mssql-server-mongodb-elasticsearch-and-others/>
- [48] G. Roden. "Datenbanktypen im Vergleich." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.heise.de/blog/Datenbanktypen-im-Vergleich-5061042.html>
- [49] AlleAktien. "objektorientierte Datenbanken." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.alleaktien.com/lexikon/objektorientierte-datenbanken>
- [50] chrissikraus. "Was ist eine hierarchische Datenbank?: Definition „Hierarchische Datenbank“." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.dev-insider.de/was-ist-eine-hierarchische-datenbank-a-07470ced107e81723ca72bee9109a615/>
- [51] Fivetran. "Database backup: Methods and benefits." Accessed: Sep. 18, 2024. [Online]. Available: <https://www.fivetran.com/learn/database-backup>
- [52] W. Khan, T. Kumar, C. Zhang, K. Raj, A. M. Roy, and B. Luo, "SQL and NoSQL Database Software Architecture Performance Analysis and Assessments—A Systematic Literature Review," *Big Data Cogn. Comput.*, vol. 7, no. 2, 2023, Art. no. 97, doi: 10.3390/bdcc7020097.
- [53] C. Sauer, H. Eichelberger, A. S. Ahmadian, A. Dewes, and J. Jürjens, "Current Industry 4.0 Platforms – An Overview: IIP-Ecosphere Whitepaper," doi: 10.5281/zenodo.4485756.
- [54] J.-S. Ok, S.-D. Kwon, C.-E. Heo, and Y.-K. Suh, "A Survey of Industrial Internet of Things Platforms for Establishing Centralized Data-Acquisition Middleware: Categorization, Experiment, and Challenges," *Sci. Program.*, vol. 2021, no. 1, 2021, Art. no. 6641562.
- [55] N. L. Rane, "ChatGPT and similar generative artificial intelligence (AI) for smart industry: role, challenges, and opportunities for Industry 4.0, Industry 5.0, and Society 5.0," *Innov. Bus. Strateg. Manag.*, pp. 10–17, 2024, doi: 10.61577/ibsm.2024.100002.
- [56] Fraunhofer. "Welcome to ICNAP Explorer." Accessed: Oct. 25, 2024. [Online]. Available: [https://icnap\\_digital\\_services.web.fec.ipt.fraunhofer.de/](https://icnap_digital_services.web.fec.ipt.fraunhofer.de/)
- [57] S. Feuerriegel, J. Hartmann, C. Janiesch, and P. Zschech, "Generative AI," *Bus. Inf. Syst. Eng.*, vol. 66, no. 1, pp. 111–126, 2024, doi: 10.1007/s12599-023-00834-7.
- [58] OpenAI. "DALL-E 2." Accessed: Oct. 10, 2024. [Online]. Available: <https://openai.com/index/dall-e-2/>
- [59] OpenAI. "GPT-4 is OpenAI's most advanced system, producing safer and more useful responses." Accessed: Oct. 10, 2024. [Online]. Available: <https://openai.com/index/gpt-4/>

- [60] Siemens. "Das Potenzial generativer KI in der Industrie nutzen." Accessed: Oct. 10, 2024. [Online]. Available: <https://www.siemens.com/de/de/produkte/automatisierung/themenfelder/tia/future-topics/industrial-copilot.html>
- [61] I. J. Goodfellow et al., "Generative Adversarial Networks," 2014, doi: 10.48550/arXiv.1406.2661.
- [62] A. Vaswani et al., "Attention Is All You Need," 2023, doi: 10.48550/arXiv.1706.03762.
- [63] L. Girin, S. Leglaive, X. Bie, J. Diard, T. Hueber, and X. Alameda-Pineda, "Dynamical Variational Autoencoders: A Comprehensive Review," *Found. Trends Mach. Learn.*, vol. 15, 1-2, pp. 1–175, 2021, doi: 10.1561/22000000089.
- [64] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," 2018. Accessed: Oct. 28, 2024. [Online]. Available: [https://cdn.openai.com/research-covers/language-unsupervised/language\\_understanding\\_paper.pdf](https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf)
- [65] T. B. Brown et al., "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems 33: 34th Conference on Neural Information Processing Systems (NeurIPS 2020). Online. 6 – 12 December 2020. Volume 1 of 27*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., Red Hook, NY: Curran, 2020, pp. 1877–1901.
- [66] C. Raffel et al., "Exploring the limits of transfer learning with a unified text-to-text transformer," *J. Mach. Learn. Res.*, vol. 21, no. 1, 2019, Art. no. 140, doi: 10.5555/3455716.3455856.
- [67] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language Models are Unsupervised Multitask Learners," Accessed: Oct. 10, 2024. [Online]. Available: [https://cdn.openai.com/better-language-models/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf)
- [68] C. Sheridan and M. Breunig. "Five use cases for manufacturers to get started with generative AI." Accessed: Oct. 28, 2024. [Online]. Available: <https://cloud.google.com/blog/topics/manufacturing/five-generative-ai-use-cases-for-manufacturing?hl=en>
- [69] G. Srinivasan, R. Gupta, N. Mittal, R. Nanda, C. Perricos, and K. Nuttal. "A new frontier in artificial intelligence: Implications of Generative AI for businesses." Accessed: Oct. 10, 2024. [Online]. Available: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-generative-artificial-intelligence.pdf>
- [70] M. Chui et al. "The economic potential of generative AI: The next productivity frontier.: The next productivity frontier." Accessed: Oct. 10, 2024. [Online]. Available: <https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/the%20economic%20potential%20of%20generative%20ai%20the%20next%20productivity%20frontier/the-economic-potential-of-generative-ai-the-next-productivity-frontier.pdf>
- [71] R. Giovis and E. Rozsa. "Transforming customer service: How generative AI is changing the game." Accessed: Oct. 10, 2024. [Online]. Available: <https://www.ibm.com/think/topics/generative-ai-for-customer-service>
- [72] P. Khetarpal. "Generative AI in Customer Service: A Game Changer for the Industry." Accessed: Oct. 10, 2024. [Online]. Available: <https://www.searchunify.com/blog/generative-ai-in-customer-service-a-game-changer-for-the-industry/>
- [73] S. Bamberger, N. Clark, S. Ramachandran, and V. Sokolova. "How Generative AI Is Already Transforming Customer Service." Accessed: Oct. 10, 2024. [Online]. Available: <https://www.bcg.com/publications/2023/how-generative-ai-transforms-customer-service>
- [74] Deloitte AI Institute. "The Generative AI Dossier: A selection of high-impact use cases across six major industries." Accessed: Oct. 10, 2024. [Online]. Available: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-gen-ai-use-cases.pdf>
- [75] A. Rajbhoj, A. Somase, P. Kulkarni, and V. Kulkarni, "Accelerating Software Development Using Generative AI: ChatGPT Case Study," in *Proceedings of the 17th Innovations in Software Engineering Conference (Formerly known as India Software Engineering Conference): February 22-24, 2024. IIITB Bangalore, India*, S. K. Chakrabarti et al., Eds., 2024, doi: 10.1145/3641399.3641403.
- [76] M. Wu, H. Brandhorst, M.-C. Marinescu, J. M. Lopez, M. Hlava, and J. Busch, "Automated metadata annotation: What is and is not possible with machine learning," *Data Intell.*, vol. 5, no. 1, pp. 122–138, 2023, doi: 10.1162/dint\_a\_00162.
- [77] J. Čelić et al., "Generative AI in E-maintenance: Myth or Reality?," in 2024 47th MIPRO ICT and Electronics Convention (MIPRO): May 20 - 24, 2024. *Opatija, Croatia. Proceedings*, S. Babic et al., Eds., 2024, pp. 1911–1919, doi: 10.1109/MIPRO60963.2024.10569282.
- [78] J. D. Brüns and M. Meißner, "Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity," *J. Retail. Consum. Serv.*, vol. 79, 2024, Art. no. 103790, doi: 10.1016/j.jretconser.2024.103790.
- [79] B. S. Haney, "AI Patents: A Data Driven Approach," *Chicago-Kent Journal of Intellectual Property*, vol. 19, no. 3, 2020, Art. no. 6. doi: 10.2139/ssrn.3527154. [Online]. Available: <https://scholarship.kentlaw.iit.edu/ckjip/vol19/iss3/6>

- [80] W. Baker et al., "Classification of Twitter Vaping Discourse Using BERTweet: Comparative Deep Learning Study," *MIR Med. Inform.*, vol. 10, no. 7, 2022, Art. no. e33678, doi: 10.2196/33678.
- [81] S. Bubeck et al., "Sparks of Artificial General Intelligence: Early experiments with GPT-4," 2023, doi: 10.48550/arXiv.2303.12712.
- [82] N. Bansal, A. Sharma, and R. K. Singh, "Fuzzy AHP approach for legal judgement summarization," *J. Manag. Anal.*, vol. 6, no. 3, pp. 323–340, 2019, doi: 10.1080/23270012.2019.1655672.
- [83] United Nations Human Rights Office of the High Commissioner. "Advancing responsible development and deployment of GenAI: The value proposition of the UN Guiding Principles on Business and Human Rights. A UN B-Tech Foundational Paper." Accessed: Oct. 28, 2024. [Online]. Available: <https://www.ohchr.org/sites/default/files/documents/issues/business/b-tech/advancing-responsible-development-and-deployment-of-GenAI.pdf>
- [84] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?," in *FACCT '21. Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency: March 3-10, 2021. Virtual Event, Canada, 2021*, pp. 610–623, doi: 10.1145/3442188.3445922.
- [85] S. Bond-Taylor, A. Leach, Y. Long, and C. G. Willcocks, "Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models," *IEEE Trans. Pattern Anal. Mach.*, vol. 44, no. 11, pp. 7327–7347, 2022, doi: 10.1109/TPAMI.2021.3116668.
- [86] G. Batra, A. Queirolo, and N. Santhanam. "Artificial intelligence: The time to act is now." Accessed: Oct. 10, 2024. [Online]. Available: <https://www.mckinsey.com/industries/industrials-and-electronics/our-insights/artificial-intelligence-the-time-to-act-is-now>
- [87] A. Urlana, C. V. Kumar, A. K. Singh, B. M. Garlapati, S. R. Chalamala, and R. Mishra, "LLMs with Industrial Lens: Deciphering the Challenges and Prospects – A Survey," 2024, doi: 10.48550/ARXIV.2402.14558.
- [88] M. U. Hadi et al., "Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects," doi: 10.36227/techrxiv.23589741.v7.
- [89] P. Leitão, "Agent-based distributed manufacturing control: A state-of-the-art survey," *Eng. Appl. Artif. Intell.*, vol. 22, no. 7, pp. 979–991, 2009, doi: 10.1016/j.engappai.2008.09.005.
- [90] E. Hofmann and M. Rüsçh, "Industry 4.0 and the current status as well as future prospects on logistics," *Comput. Ind.*, vol. 89, pp. 23–34, 2017, doi: 10.1016/j.compind.2017.04.002.
- [91] H. Kagermann, W. Wahlster, and J. Helbig. "Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry. Final report of the Industrie 4.0 Working Group." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.din.de/resource/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf>
- [92] P. J. Gisi, *The Dark Factory and the Future of Manufacturing: A Guide to Operational Efficiency and Competitiveness*. New York, NY: Routledge, 2024.
- [93] P. Osterrieder, L. Budde, and T. Friedli, "The smart factory as a key construct of Industry 4.0: A systematic literature review," *Int. J. Prod. Econ.*, vol. 221, 2020, Art. no. 107476, doi: 10.1016/j.ijpe.2019.08.011.
- [94] Q. Gong, G. Chen, W. Zhang, and H. Wang, "The role of humans in flexible smart factories," *Int. J. Prod. Econ.*, vol. 254, 2022, Art. no. 108639, doi: 10.1016/j.ijpe.2022.108639.
- [95] E. Oztemel and S. Gursev, "Literature review of Industry 4.0 and related technologies," *J. Intell. Manuf.*, vol. 31, no. 1, pp. 127–182, 2020, doi: 10.1007/s10845-018-1433-8.
- [96] P. Osterrieder, L. Budde, and T. Friedli, "The smart factory as a key construct of Industry 4.0: A systematic literature review," *Int. J. Prod. Econ.*, vol. 221, 2020, Art. no. 107476, doi: 10.1016/j.ijpe.2019.08.011.
- [97] Q. Gong, G. Chen, W. Zhang, and H. Wang, "The role of humans in flexible smart factories," *Int. J. Prod. Econ.*, vol. 254, 2022, Art. no. 108639, doi: 10.1016/j.ijpe.2022.108639.
- [98] S. Boom. "Remote commissioning: why you need it and how to use it." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.ixon.cloud/knowledge-hub/remote-commissioning>
- [99] genua. "Remote Maintenance: Customized Adaptation to Individual Remote Maintenance Situations." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.genua.eu/it-security-solutions/remote-maintenance>
- [100] W. Krenz and D. Kronenwett. "Is "Pay-Per-Use" The Future In Manufacturing Industries?: An innovative business model may not live up to the expectations." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.oliverwyman.com/our-expertise/insights/2019/nov/perspectives-on-manufacturing-industries-vol-14/manufacturing-in-a-changing-world/is-pay-per-use-the-future-in-manufacturing-industries.html>
- [101] TRUMPF. "Pay per Part: Mehr produzieren ohne investieren." Accessed: Sep. 12, 2024. [Online]. Available: [https://www.trumpf.com/de\\_DE/produkte/services/services-maschinen-systeme-und-laser/pay-per-part/](https://www.trumpf.com/de_DE/produkte/services/services-maschinen-systeme-und-laser/pay-per-part/)



- [102] M. Schneider, F. Haag, A. K. Khalil, and D. A. Breunig, "Evaluation of Communication Technologies for Distributed Industrial Control Systems: Concept and Evaluation of 5G and WiFi 6," *Procedia CIRP*, vol. 107, pp. 588–593, 2022, doi: 10.1016/j.procir.2022.05.030.
- [103] Dimitri-furman *et al.* "Hyperscale service tier." Accessed: Dec. 9, 2024. [Online]. Available: <https://learn.microsoft.com/en-us/azure/azure-sql/database/service-tier-hyperscale?view=azuresql>
- [104] Siemens. "Predictive Maintenance – Zuverlässigkeit neu definiert." Accessed: Sep. 12, 2024. [Online]. Available: [https://www.siemens.com/de/de/produkte/services/digital-enterprise-services/analytik-kuenstliche-intelligenz-services/predictive-services.html?gclid=EAlaIqobChMlle-RmYu9iAMVP4ODBx21IBZQEAAAYASAAEglpZfD\\_BwE&acz=1&gad\\_source=1](https://www.siemens.com/de/de/produkte/services/digital-enterprise-services/analytik-kuenstliche-intelligenz-services/predictive-services.html?gclid=EAlaIqobChMlle-RmYu9iAMVP4ODBx21IBZQEAAAYASAAEglpZfD_BwE&acz=1&gad_source=1)
- [105] R. Kaviyaraj and M. Uma, "A Survey on Future of Augmented Reality with AI in Education," in *International Conference on Artificial Intelligence and Smart Systems (ICAIS 2021): 25-27, March 2021. JCT College of Engineering and Technology Coimbatore, India. Proceedings, 2021*, pp. 47–52, doi: 10.1109/ICAIS50930.2021.9395838.
- [106] S. Choi, K. Jung, and S. D. Noh, "Virtual reality applications in manufacturing industries: Past research, present findings, and future directions," *Concurrent Engineering*, vol. 23, no. 1, pp. 40–63, 2015, doi: 10.1177/1063293X14568814.
- [107] M. Jacoby, F. Volz, C. Weißenbacher, and J. Müller, "FA<sup>3</sup>ST Service – An Open Source Implementation of the Reactive Asset Administration Shell," in *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2022, doi: 10.1109/ETFA52439.2022.9921584.
- [108] S. Choi, K. Jung, and S. D. Noh, "Virtual reality applications in manufacturing industries: Past research, present findings, and future directions," *Concurrent Engineering*, vol. 23, no. 1, pp. 40–63, 2015, doi: 10.1177/1063293X14568814.
- [109] MachineMetrics. "Industrial Communication Protocols." Accessed: Sep. 16, 2024. [Online]. Available: <https://www.machinemetrics.com/connectivity/protocols>
- [110] D. Gamero, A. Dugenske, C. Saldana, T. Kurfess, and K. Fu, "Scalability Testing Approach for Internet of Things for Manufacturing SQL and NoSQL Database Latency and Throughput," *J. Comput. Inf. Sci. Eng.*, vol. 22, no. 6, 2022, Art. no. 060901, doi: 10.1115/1.4055733.
- [111] Microsoft. "The best ideas need the best AI platform." Accessed: Sep. 12, 2024. [Online]. Available: <https://azure.microsoft.com/en-us>
- [112] Siemens. "Industrial Edge – maximize your competitive edge." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.siemens.com/global/en/products/automation/topic-areas/industrial-edge.html>
- [113] Broadcom. "Hybrid Cloud Solutions." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.vmware.com/solutions/cloud-infrastructure/hybrid-cloud>
- [114] Docker Inc. "Develop faster. Run anywhere.: Build with the #1 most-used developer tool." Accessed: Oct. 9, 2024. [Online]. Available: <https://www.docker.com/>
- [115] Global Services Security Team, Amazon Web Services. "AWS Security Reference Architecture (AWS SRA)." Accessed: Sep. 9, 2024. [Online]. Available: <https://docs.aws.amazon.com/prescriptive-guidance/latest/security-reference-architecture/welcome.html>
- [116] Microsoft. "Azure Data Encryption at rest." Accessed: Sep. 16, 2024. [Online]. Available: <https://learn.microsoft.com/en-us/azure/security/fundamentals/encryption-atrest>
- [117] msmbaldwin *et al.* "Azure encryption overview." Accessed: Sep. 16, 2024. [Online]. Available: <https://learn.microsoft.com/en-us/azure/security/fundamentals/encryption-overview#encryption-of-data-in-transit>
- [118] ML6. "AI Machine Vision for manufacturers: ML6 Solution for manufacturers." Accessed: Sep. 16, 2024. [Online]. Available: <https://www.ml6.eu/domains/computer-vision/ai-machine-vision>
- [119] N. Marz and W. James, *Big data: Principles and best practices of scalable real-time data systems*. Shelter Island, NY: Manning, 2015.
- [120] ABB. "Industrial IoT applications: How ABB creates value from the Industrial Internet of Things." Accessed: Sep. 8, 2024. [Online]. Available: <https://new.abb.com/control-systems/features/industrial-iiot-services-people-use-cases>
- [121] MDPI. "Environmental Sensing: A section of Sensors (ISSN 1424-8220)." Accessed: Aug. 9, 2024. [Online]. Available: [https://www.mdpi.com/journal/sensors/sections/environmental\\_sensing](https://www.mdpi.com/journal/sensors/sections/environmental_sensing)
- [122] SRI International. "Computational sensing and embedded low-power processing." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.sri.com/research/information-computing-sciences/computer-vision/computational-sensing-embedded-low-power-processing/#:~:text=Computational%20sensing%20is%20where%20optics,information%20for%20a%20given%20application>
- [123] P. Denny *et al.*, "Generative AI for Education (GAIED): Advances, Opportunities, and Challenges," 2024, doi: 10.48550/arXiv.2402.01580.
- [124] P. Alexopoulos, *Semantic Modeling for Data: Avoiding Pitfalls and Breaking Dilemmas*. Beijing, Boston, Farnham, Sebastopol, Tokyo: O'reilly, 2020.
- [125] OPC Foundation. "UA Companion Specifications." Accessed: Sep. 12, 2024. [Online]. Available: <https://opcfoundation.org/about/opc-technologies/opc-ua/ua-companion-specifications/>

- [126] NI Apps. "The right sample rate on your DAQ device makes all the difference - DAQ Fundamentals." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.youtube.com/watch?v=naau6Nyf9Qc>
- [127] CHROMOS Industrial. "Einrichtung eines Cobot von Universal Robots | Vorteile von Cobots | CHROMOS Industrial." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.youtube.com/watch?v=xWyku8N2aGA>
- [128] A. Dhaliwal, "The Rise of Automation and Robotics in Warehouse Management," in *Transforming Management Using Artificial Intelligence Techniques* (Artificial intelligence (AI). Elementary to advanced practices), V. Garg and R. Agrawal, Eds., Boca Raton: CRC Press, 2020, pp. 63–72.
- [129] X. Zhao and T. Chidambareswaran, "Autonomous Mobile Robots in Manufacturing Operations," in *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE): August 26-30, 2023, Auckland, New Zealand*, 2023, doi: 10.1109/CASE56687.2023.10260631.
- [130] Tech Vision. "Inside Amazon's Smart Warehouse." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.youtube.com/watch?v=IMPbKVb8y8s>
- [131] matellio. "Unveiling the Power of M2M in Manufacturing." Accessed: Dec. 9, 2024. [Online]. Available: <https://www.matellio.com/blog/m2m-in-manufacturing/>
- [132] M. Jafari et al., "Emotion recognition in EEG signals using deep learning methods: A review," *Comput. Biol. Med.*, vol. 165, 2023, Art. no. 107450, doi: 10.1016/j.compbimed.2023.107450.
- [133] Fraunhofer Institute for Production Technology IPT. "Research project "CAMStylus"" Accessed: Sep. 12, 2024. [Online]. Available: <https://www.ipt.fraunhofer.de/en/projects/camstylus.html>
- [134] Trilateral Research. "Ethical AI Pillars: Human Agency and Oversight." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.youtube.com/watch?v=LFbFJUZCINs>
- [135] StartUs Insights. "5 Top Emerging Manufacturing-as-a-Service (MaaS) Startups." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.startus-insights.com/innovators-guide/5-top-emerging-manufacturing-as-a-service-maas-startups/>
- [136] O. Fisher, N. Watson, L. Porcu, D. Bacon, M. Rigley, and R. L. Gomes, "Cloud manufacturing as a sustainable process manufacturing route," *J. Manuf. Syst.*, vol. 47, pp. 53–68, 2018, doi: 10.1016/j.jmsy.2018.03.005.
- [137] D. Sanchez-Londono, G. Barbieri, and L. Fumagalli, "Smart retrofitting in maintenance: a systematic literature review," *J. Intell. Manuf.*, vol. 34, no. 1, pp. 1–19, 2023, doi: 10.1007/s10845-022-02002-2.
- [138] T. H. Khan, C. Noh, and S. Han, "Correspondence measure: a review for the digital twin standardization," *Int. J. Adv. Manuf. Technol.*, vol. 128, 5-6, pp. 1907–1927, 2023, doi: 10.1007/s00170-023-12019-3.
- [139] Fraunhofer Institute for Production Technology IPT. "Research project "CAM2030"" Accessed: Dec. 9, 2024. [Online]. Available: <https://www.ipt.fraunhofer.de/en/projects/cam2030.html>
- [140] X. Shao, C. You, and R. Zhang, "Intelligent Reflecting Surface Aided Wireless Sensing: Applications and Design Issues," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 383–389, 2024, doi: 10.1109/MWC.004.2300058.
- [141] K. Y. H. Lim, P. Zheng, and C.-H. Chen, "A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives," *J. Intell. Manuf.*, vol. 31, no. 6, pp. 1313–1337, 2020, doi: 10.1007/s10845-019-01512-w.
- [142] J. Knußmann and R. H. Schmitt, "Real-Time Digital Twin," in *ICNAP Study Report 2022: International Center for Networked, Adaptive Production, Aachen*, 2023. Accessed: Oct. 23, 2024. [Online]. Available: <https://publica.fraunhofer.de/entities/publication/72730edf-bad3-47b0-885b-9f01654051ab/details>
- [143] R. A. Atmoko, R. Riantini, and M. K. Hasin, "IoT real time data acquisition using MQTT protocol," *J. Phys.: Conf. Ser.*, vol. 853, 2017, Art. no. 012003, doi: 10.1088/1742-6596/853/1/012003.
- [144] The OnQ Team. "We are making AI ubiquitous." Accessed: Sep. 9, 2024. [Online]. Available: <https://www.qualcomm.com/news/onq/2020/06/we-are-making-ai-ubiquitous>
- [145] T. Do-Duy, D. van Huynh, O. A. Dobre, B. Canberk, and T. Q. Duong, "Digital Twin-Aided Intelligent Offloading With Edge Selection in Mobile Edge Computing," *IEEE Wireless Commun. Lett.*, vol. 11, no. 4, pp. 806–810, 2022, doi: 10.1109/LWC.2022.3146207.
- [146] A. Z. Faroukhi, I. El Alaoui, Y. Gahi, and A. Amine, "Big data monetization throughout Big Data Value Chain: a comprehensive review," *J. Big Data*, vol. 7, no. 1, 2020, Art. no. 3, doi: 10.1186/s40537-019-0281-5.
- [147] German Federal Ministry for Economic Affairs and Climate Action. "The Gaia-X Ecosystem - A Sovereign Data Infrastructure for Europe." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.bmwk.de/Redaktion/EN/Dossier/gaia-x.html>
- [148] Z. Huang, Y. Shen, J. Li, M. Fey, and C. Brecher, "A Survey on AI-Driven Digital Twins in Industry 4.0: Smart Manufacturing and Advanced Robotics," *Sensors*, vol. 21, no. 19, 2021, Art. no. 6340, doi: 10.3390/s21196340.
- [149] Fraunhofer Institute for Integrated Circuits IIS. "Integrated AI for Sensor Systems." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.iis.fraunhofer.de/en/ff/sse/sensor-solutions/integrated-ai-for-sensor-systems.html>

- [150] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB. "Understand and comprehend AI models." Accessed: Sep. 12, 2024. [Online]. Available: [https://www.iosb.fraunhofer.de/en/competences/image-exploitation/human-ai-interaction/explainable-ai.html#copy\\_1434514291](https://www.iosb.fraunhofer.de/en/competences/image-exploitation/human-ai-interaction/explainable-ai.html#copy_1434514291)
- [151] T. M. Anandan. "Plug-and-Play Robot Ecosystems on the Rise." Accessed: Sep. 12, 2024. [Online]. Available: <https://www.automate.org/robotics/industry-insights/plugin-and-play-robot-ecosystems-on-the-rise>
- [152] J. Wan, S. Tang, H. Yan, D. Li, S. Wang, and A. V. Vasilakos, "Cloud Robotics: Current Status and Open Issues," *IEEE Access*, vol. 4, pp. 2797–2807, 2016, doi: 10.1109/ACCESS.2016.2574979.
- [153] S. N. Raisin, J. Jamaludin, F. Mohd Rahalim, F. A. Jamal Mohamad, and B. Naeem, "Cyber-Physical System (CPS) Application- A REVIEW," *REKA ELKOMIKA*, vol. 1, no. 2, pp. 52–65, 2020, doi: 10.26760/rekaelkomika.v1i2.52-65.
- [154] S. Wang, J. Wan, D. Li, C. Zhang, and Di Li, "Implementing Smart Factory of Industrie 4.0: An Outlook," *Int. J. Distrib. Sens. Netw.*, 2016, Art. no. 3159805, doi: 10.1155/2016/3159805.
- [155] S. Wang, J. Wan, D. Zhang, L. Li, and C. Zhang, "Towards smart factory for Industry 4.0: a self-organized multi-agent system with big data based feedback and coordination," *Comput. Netw.*, vol. 101, pp. 158–168, 2016, doi: 10.1016/j.comnet.2015.12.017.
- [156] C. Chen, T. L. Kong, and W. Kan, "Identifying the promising production planning and scheduling method for manufacturing in Industry 4.0: a literature review," *Prod. Manuf. Res.*, vol. 11, no. 1, 2023, Art. no. 2279329, doi: 10.1080/21693277.2023.2279329.
- [157] S. Wang, J. Wan, D. Zhang, L. Li, and C. Zhang, "Towards smart factory for Industry 4.0: a self-organized multi-agent system with big data based feedback and coordination," *Comput. Netw.*, vol. 101, pp. 158–168, 2016, doi: 10.1016/j.comnet.2015.12.017.
- [158] C. Chen, T. L. Kong, and W. Kan, "Identifying the promising production planning and scheduling method for manufacturing in Industry 4.0: a literature review," *Prod. Manuf. Res.*, vol. 11, no. 1, 2023, Art. no. 2279329, doi: 10.1080/21693277.2023.2279329.
- [159] S. Z. Kamal, S. M. Al Mubarak, B. D. Scodova, P. Naik, P. Flichy, and G. Coffin, "IT and OT Convergence - Opportunities and Challenges," in *SPE Intelligent Energy International Conference and Exhibition 2016: Aberdeen, United Kingdom, 6-8 September 2016*, 2016, No. SPE-181087-MS, doi: 10.2118/181087-MS.
- [160] W. Maass et al., "Quantum Computing Enhanced Service Ecosystem for Simulation in Manufacturing," 2024, doi: 10.48550/arXiv.2401.10623.
- [161] S. Z. Kamal, S. M. Al Mubarak, B. D. Scodova, P. Naik, P. Flichy, and G. Coffin, "IT and OT Convergence - Opportunities and Challenges," in *SPE Intelligent Energy International Conference and Exhibition 2016: Aberdeen, United Kingdom, 6-8 September 2016*, 2016, No. SPE-181087-MS, doi: 10.2118/181087-MS.
- [162] W. Maass et al., "Quantum Computing Enhanced Service Ecosystem for Simulation in Manufacturing," 2024, doi: 10.48550/arXiv.2401.10623.
- [163] V. Bhargava and R. Wagner. "Next big thing in smart factories? Control systems virtualization." Accessed: Sep. 9, 2024. [Online]. Available: <https://iebmedia.com/technology/industrial-ethernet/the-next-big-thing-in-smart-manufacturing-control-systems-virtualization/>
- [165] E. Gent. "Silicon Valley Is Reviving the Dream of General-Purpose Humanoid Robots." Accessed: Sep. 12, 2024. [Online]. Available: <https://singularityhub.com/2023/05/22/silicon-valley-is-reviving-the-dream-of-general-purpose-humanoid-robots/>

## Authors

---

Alexander D. Kies, Mario Pothen,  
Henrik Heymann, Niels König,  
Max Ortmann, Robert H. Schmitt,  
André Gilerson, Alexander Mattern,  
Jan Hendrik Hellmich,  
Ines Groß-Weege, Liz Leutner,  
David Wichter, Thomas Bergs,  
Alexander Frigge, Janina Gauß,  
Tim Geerken, Andrea Lanfermann,  
Bastian Thanhäuser

## Contact

---

Alexander Kies  
ICNAP Community Manager  
Phone +49 241 8904-498  
[community@icnap.de](mailto:community@icnap.de)

International Center for Networked,  
Adaptive Production  
c/o Fraunhofer Institute for  
Production Technology IPT  
Steinbachstrasse 17  
52074 Aachen  
Germany  
[www.icnap.de](http://www.icnap.de)

DOI: 10.24406/publica-3748