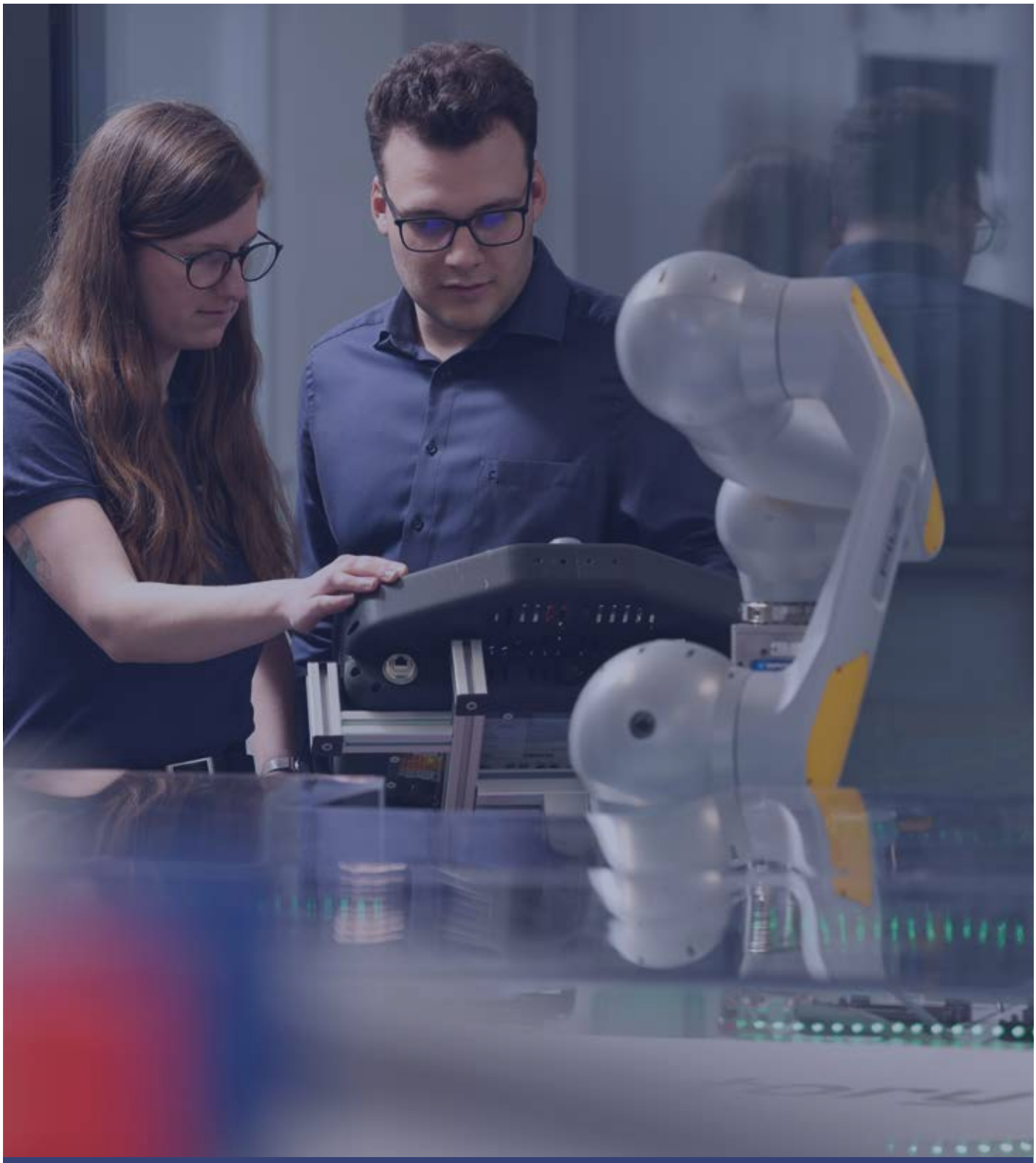

Trustworthiness: Safety Meets Security in Industry 4.0 Ecosystems

Whitepaper SF-3.5: 04/2024

smartFactory KL®

Abstract
Current challenges facing the manufacturing sector necessitate innovations such as modular production. This novel form of production can be dynamically aligned with market demand and provides a highly efficient, hence sustainable, production environment. A key element for flexibility and efficiency is data exchange between production machines and the company's IT systems. As recent cyberattacks demon-strate, more communication results in more cyber risk which must be mitigated by effective cybersecuri-ty. This paper proposes a holistic view of machine safety and cybersecurity called trustworthiness. The paper then intro-duces the concept of predictive trustworthiness, which allows the potential conse-quences of cyber or other runtime incidents to be estimated by using information from digital twins in conjunction with knowledge graphs. This enables autonomous or enhanced operator-controlled re-sponses that increase productivity while remaining safe. Finally, the applicability of the concept is demonstrated in a use-case scenario at the modular production system of ***SmartFactory***^{KL}.

Keywords
Trustworthiness, Functional Safety, Cybersecurity, Resilience, Reliability, Privacy, Productivity, Integrity, Plausibility, Industry 4.0, Modular Production, Digital Twin, Asset Administration Shell

Authors:

Alexander Kurdas	TÜV SÜD Industrie Service GmbH
Michael Pfeifer	TÜV SÜD Industrie Service GmbH
Dr. Detlev Richter	TÜV SÜD Product Service GmbH
Ben-Frederik Ehlers	Fujitsu
Jamie Wilkie	Fujitsu
Stefan Meny	Fujitsu
Bernd Neuschwander	Pilz GmbH & Co. KG
Bernd Eisenhuth	Pilz GmbH & Co. KG
Marco Sprenger	B&R
Jonathan Nußbaum	RPTU Kaiserslautern-Landau
Dr. Henning Gösling	DFKI GmbH
Philipp Richard	DFKI GmbH / Technologie-Initiative ***SmartFactory***^{KL} e.V./ RPTU
Prof. Dr.-Ing. Martin Ruskowski	DFKI GmbH / Technologie-Initiative ***SmartFactory***^{KL} e.V./ RPTU

Content

1. Motivation

In the current fast changing world, manufacturing companies are faced with major challenges, such as turbulences in supply chains and more individual customer wishes coupled with an increasing variety of products. This leads to smaller batch sizes and ever shorter product life cycles. To master these challenges economically, machines and factories need to become more flexible and agile. Communication and data play an ever greater role in the management, improvement, and control of production flexibility, efficiency, and sustainability. Operators recognize that reaching their business objectives requires greater exploitation of data from their own production processes. Production systems, that used to be operated offline, are increasingly connected to local Information Technology (IT), the internet or even cloud applications. But with networking comes a risk: cyber-attacks. This includes malware attacks or industrial espionage but also manipulation of Operational Technology (OT) systems. Especially misuse or deactivation of safety functions can endanger people and the environment. Conversely, a safety function can also be used to stop production and thus impact productivity. To protect companies against such attacks, cybersecurity is well-established in IT systems while being typically much less established in OT systems. Hence, this paper offers a bridge between these disciplines.

The gap between IT and OT results from different underlying priorities, which lead to conflicts and contradictions when linking both. While IT sees data security as the top priority, OT focuses on productivity and safety. For example, IT tries to keep software up to date to fix security issues, whereas OT avoids updates to limit negative influences on running production. Especially functional safety updates are problematic as costly re-certifications are required. This shows that the current underlying design idea of safety is not directed towards a fast changing and flexible operation of the systems. Hence connections to the ever-evolving internet are problematic for safety and need special attention. However, operators often lack IT and, especially, cybersecurity skills. This leads to uncertainty about appropriate measures to apply and additionally it encourages cyber-attacks on “soft” but valuable targets.

Reticence against security in OT also stems from the fact that cybersecurity requirements and measures are rather recent, unfamiliar, and not yet settled within the industry. This contrasts with the well-established safety measures, which have evolved over decades, with a clear and comprehensive regulatory framework. While safety measures are enforced by law to protect workers, security measures should be applied to protect the company itself. Regulatory requirements regarding security in conjunction with safety have recently begun to evolve. For example, in Germany the Technical Rule for Operational Safety (TRBS) 1115-1 [1] of 2022 details the cybersecurity requirements in industry that must be fulfilled for safety-relevant measurement and control equipment. Furthermore, the European Parliament recently passed the new version of the machinery regulation [2], including statements considering connections between machinery and cybersecurity, among the well-known safety regulations. Additionally, more cybersecurity regulations are applicable to or are explicitly addressing aspects of OT, e.g. the NIS 2 directive of the EU [3], the NIST Cybersecurity Framework (CSF) [4] or the standards series IEC 62443 [5]. To fulfill these upcoming new regulations, it is time to embrace the convergence of IT and OT and to rethink security and safety concepts in industry.

2. Objective

Trustworthiness is the overall term that combines the topics of safety, cybersecurity, privacy, reliability, and resilience and their intersections, as depicted in Figure 1. The term stands for a holistic view of all sub-areas and is by now established in industry. For example, the Trust Vector concept presented in [6] can be used to ensure trustworthiness in real-time communication. This white paper provides an overview of how trustworthiness can be achieved in a dynamic environment with the help of digital twins, anomaly detection, and predictive safety methods. It shows how contradictions between safety and cybersecurity, such as differing latencies, can be resolved. In addition, we show how the concept of the digital twin fits into the conceptual landscape of skill-based production, operational safety intelligence, the Purdue Model, and how the digital twin connects these topics. The concept and its implementation will be demonstrated using the hazardous goods transport use case, shown live at the SmartFactory-KL exhibition stand at the Hannover Messe 2024.

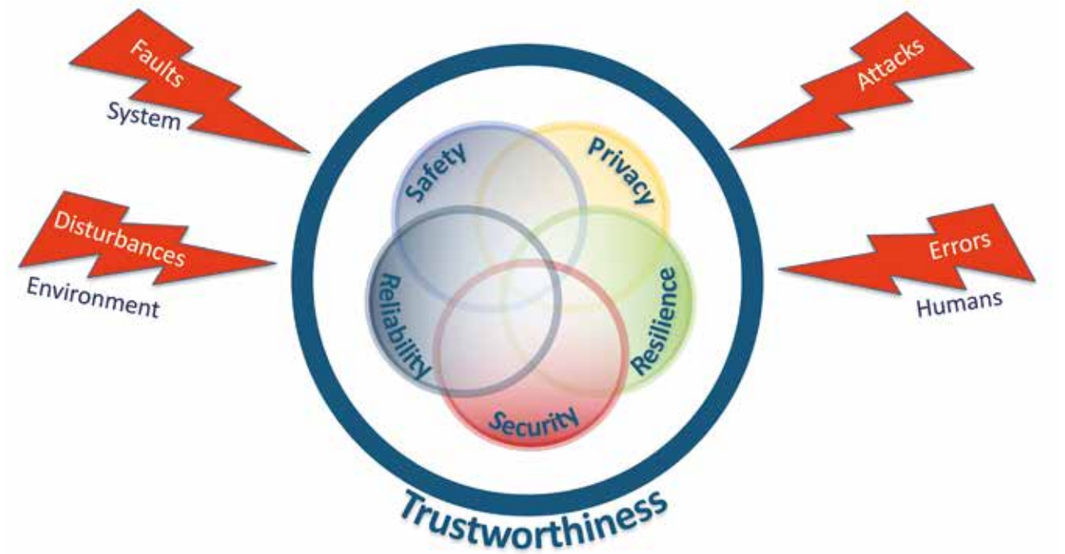

Figure 1: Trustworthiness and its influences[1]

[1] Source: Wibu-Systems AG, https://www.wibu.com/de/blog/article/permeation-of-trust-in-iiot-systems.html

3. The Production Level 4 Ecosystem

The ***Production Level 4*** ecosystem refers to a future-oriented, resilient, and flexible production method that takes an integrative view of humans and machines. It is based on a modular, subsidiary approach that encapsulates the complexity of automation technology through local intelligent control and computation systems. This approach enables flexible reconfiguration of production modules and focuses on sustainability by taking both environmental and social aspects into account. The ***Production Level 4*** vision also includes the implementation of Industry 4.0 technologies, the use of cloud and edge computing for data processing and reliable, autonomous production control. At the heart of this is the idea of "shared production", in which production services are made available on demand in a network of trusted partners [7]. The ***Production Level 4*** ecosystem requires various technologies, and one of the key enablers is the digital twin. A digital twin serves as a dynamic digital representation of a physical system or process across its lifecycle, utilized for simulation, analysis, control and other aspects as depicted in Figure 2. It supports the modular and subsidiary approach by enabling the simulation and optimization of production processes before physical changes are made, thus encapsulating the complexity of automation technology through local intelligent control and computation systems. Complex relationships and dependencies within the digital twin, such as hazard and safety management, can be managed using knowledge graphs. By structuring data and relationships in a graph format, knowledge graphs provide transparent and navigable means to analyze the safety aspects of production processes, assets and human interactions [8].

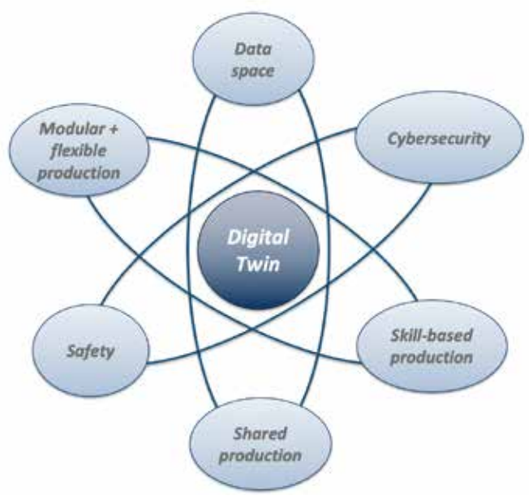

Figure 2: Digital twin subject areas drawn as an atom

The core concept of shared production is the availability of readily exchangeable software and hardware resources, enabling the reconfiguration into different systems on demand. Requiring minimal reconfiguration effort, trusting partners offer their production and manufacturing capacities and capabilities as necessary, resulting in a collaborative manufacturing network of decentralized participants. Shared production emphasizes transparency in data, process flexibility and sustainability. The system benefits manufacturers of all sizes and increases the opportunity for small and medium sized enterprises to contribute to and participate in the system [7].

Skill-based Production focuses on developing standardized, manufacturer-independent interfaces and interactions for reconfigurable and flexible manufacturing processes. By enabling seamless reconfiguration of production workflows it enhances adaptability and efficiency in production environments. Skill-based production is a key enabler for shared production [7].

Flexible and modular production refers to the ability to quickly adapt manufacturing processes to changing demands and requirements. As such, together with the skill-based production approach, it is a cornerstone of a shared production. Modular production systems allow for the reconfiguration of production modules to accommodate different products or production volumes efficiently. Flexibility in production enables companies to respond swiftly to market changes and customer needs. The approach involves holistically designing production systems at a hardware and software level to be interchangeable and easy to communicate with. This approach enables quick adaptation to changing market demands, minimizes downtime, and enhances overall efficiency in production operations [7], [9].

Multi-agent systems (MAS) play a crucial role in the coordination and decision-making processes within complex production environments. These systems consist of autonomous agents that interact with each other to achieve common production goals. MAS enable decentralized control, adaptability and self-organization in manufacturing systems, allowing for efficient resource allocation, task allocation, and real-time decision-making. By leveraging MAS, production systems can dynamically respond to changes, optimize production processes and enhance overall system performance [7].

Cyber-physical production modules (CPPMs) integrate physical components with digital technologies to create smart and interconnected manufacturing systems. They are modular, autonomous elements in a flexible production environment that provide specific functionalities while being reconfigurable to form adaptable production systems. These modules combine sensors, actuators, and control systems with communication networks and data processing capabilities. By bridging the physical and digital realms, CPPMs interact with their surroundings in a context-aware manner, featuring standardized interfaces, interactions and self-contained functionalities, as per the skill-based approach. By combining the functionalities of multiple CPPMs, hierarchical and flexible production structures can be created, facilitating quick adjustments with minimal effort. However, managing the complexity associated with CPPMs requires hierarchical encapsulation at different production levels to ensure operational efficiency and maintainability [7], [9].

The Asset Administration Shell (AAS) serves as a digital representation of physical assets in the production environment, encapsulating all relevant information about an asset's identity, capabilities, and lifecycle. It enables seamless communication and interaction between assets, systems, and services in an Industrial Internet of Things (IIoT) ecosystem, empowering interoperability, transparency, and data exchange in smart manufacturing environments through a standardized way to describe and access asset information. It plays a pivotal role in enabling digital twins, predictive maintenance, and data-driven decision-making, ultimately optimizing production processes and driving innovation in the Industry 4.0 landscape.

4. Cybersecurity on the Shopfloor

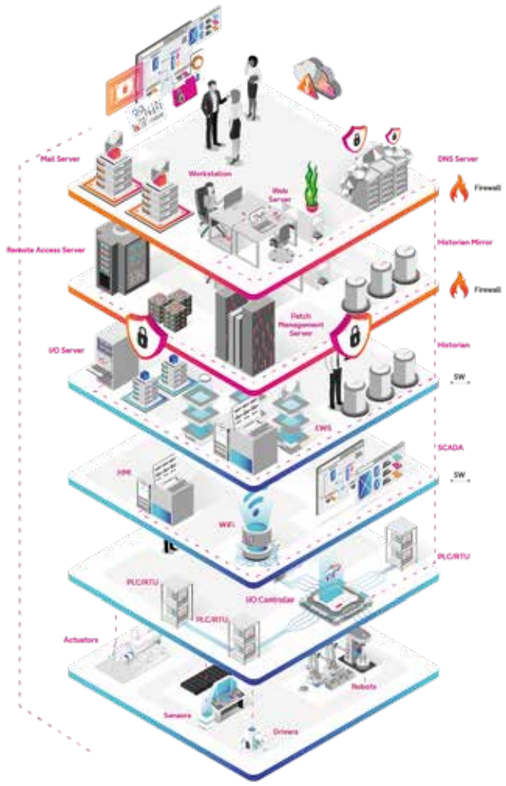

Figure 3: Purdue Model as developed in [10]

Traditionally, cybersecurity is seen as a challenge for IT systems, servers and computers used by personnel. But in the manufacturing industry cybersecurity is a multi-layer challenge for the whole company that reaches into the shopfloor and production. The Purdue Model [10] describes the generic communication and automation architecture of a company. The model as depicted in Figure 3 reads from bottom to top:

- At level 0 there are the production machines, including e.g. sensors and actuators.
- Level 1 shows specialized electronic devices which control them; Programmable Logic Controllers (PLCs) and Remote Terminal Units (RTUs) are examples.
- At level 2 multiple devices are corralled to work together, for instance on a single production line. They are controlled by Supervisory Control and Data Acquisition systems (SCADA) and are managed by Human Machine Interfaces (HMI). At this level IT devices (with Windows/Linux) are used as a platform for OT functions (SCADA/HMI), often in a frozen configuration together with the lower layers.
- At level 3 Manufacturing Execution Systems (MES) and supporting systems such as Historians coordinate the work of an entire plant.
- The blue layers 4 and 5 at the top represent the world of classic IT, e.g. backend systems such as Enterprise Resource Planning (ERP) or office workplace systems.

4.1. Challenges of Cybersecurity for OT Equipment

Attack Surface of OT Equipment
The Purdue Model [10] presented in the previous section reveals much of the OT security attack surface:

- Insufficient network segregation between level 3 and 4 may allow malware to laterally enter the OT network from IT level 5, for instance, introduced through a malicious email attachment.
- Remote access to level 2 and 3 as used for equipment maintenance may open attack paths. This happened in a well-publicized attack on the Florida water supply in 2021 .
- Level 2 devices are often in a frozen configuration with the lower levels and hence they are often outdated regarding IT and cybersecurity standards.
- The IT systems in level 2 and 3, especially the older IT systems typically found in level 2, have little inherent defense against threats like ransomware. This is currently the most common form of cyberattack, even in OT. Pilz, a member of the ***SmartFactory***^{KL} consortium, was the victim of such an attack in 2019 and has openly shared its experience and learnings .
- The controllers at level 1 are susceptible to more sophisticated attacks which deliberately manipulate the behavior of machines at level 0, as seen in an attack on an Iranian steel mill in 2022. Attackers appear to have manipulated the low-level behavior of the ladle metallurgy process, causing a fire at the plant .
- Additionally, human beings are a risk in OT as well as IT. The careless use of a USB stick in an HMI or SCADA system can introduce an infection.

[2] What's most interesting about the Florida water system hack? That we heard about it at all. – Krebs on Security
[3] https://www.pilz.com/en-US/company/news/articles/215337

Differences between IT and OT

At the next level of detail, it is important to recognize differences between IT and OT, which impact the way cybersecurity is implemented on shopfloor level. For example, there are often organizational and budget barriers between IT and OT which makes process and practice alignment hard, even though effective OT security needs to go hand in hand with IT security. This partly stems from the lack of IT and especially cyber skills in production. One consequence is that cybersecurity processes such as incident responses are frequently missing in production.

Another difference is characterized by the OT equipment in use and the capabilities of these components. The OT equipment uses industrial protocols such as Modbus, ProfiNET and EtherCAT which are not understood by pure play IT security tools. Furthermore, real-time constraints are introduced in these systems, which prohibit time-consuming investigations of network traffic. Additionally, OT networks are often quite "fragile", as they operate near to the capacity limit. Hence, they may fail if subjected to unforeseen traffic as caused by traditional IT security techniques such as active scanning of devices.

In contrast to usual IT components, OT equipment at level 0 often has a very long operational life. Decades of service are not uncommon. The controlling equipment at level 1 and 2 can be equally old and no longer receive any updates. Vulnerability patching as practiced in IT can therefore be hard or impossible to conduct for these OT systems. A further challenge for updating control devices is the need to re-certify updated functional safety systems, which is cumbersome and costly.

Finally, the detection of cyber events typically involves much higher delays than the detection of safety events. Safety systems often detect and respond to anomalies within small fractions of a second. Cyber detection may take several minutes or longer. Cyber response in OT typically requires consultation amongst experts to evaluate the impact of a response on current production. Apart from well-understood firewall rules, response is not automatically triggered.

[4] Industrial Cyber Attack on Iranian Steel Companies Explained | SCADAfence

4.2. Approaches for Cybersecurity in Production

Cybersecurity Standards

Concluding from the presented challenges, it is necessary to introduce cybersecurity controls and best practices to OT networks while respecting the special operational conditions in production. An important cybersecurity guideline which is applicable in both IT and OT is the NIST cybersecurity Framework (CSF) [4]. The most recent version 2.0 from 2024 is represented at a high level in Figure 4. In a continuously managed cycle, organizations determine their cyber exposure and risk tolerance and implement appropriate organizational, procedural and technical measures to protect against cyber risks. In this way, they are prepared for the detection of cyber incidents that do occur and can then respond to and recover from incidents. Further technical methods for cybersecurity are described in standards such as IEC 62443 [5] or NIST SP 800-82 Rev. 3 [11]. They build on established cybersecurity standards such as ISO/IEC 27001 [12] but address the needs of production. The implementation of these standards and appropriate best practices is increasingly mandated by legislation. In the European Union the NIS 2 Directive [3] addresses IT and OT security of the production process while the Machinery Regulation [2] addresses the cyber robustness of new production machinery.

Figure 4: Guideline of NIST CSF on a high-level [4]

Cybersecurity Methods
Based on the standards established best practices to increase cybersecurity in production include:

- **IT/OT segmentation:** Maintaining a controlled interface between corporate and production networks allows controlled data exchange but reduces the likelihood of malware moving laterally from the IT network into OT.

- **Micro-segmentation:** Protective mechanisms such as micro-segmentation can protect unpatchable production equipment.

- **OT-aware firewalls:** Certain firewalls are able to detect and block known malware which targets low-level devices at level 1. The detection patterns are constantly updated.

- **Secure remote access:** Available solutions control access to critical production resources.

- **Policies and processes:** Cybersecurity policies and processes which are familiar in IT can be extended and adapted for OT. An example is an incident response process which defines roles and responsibilities if a cyber incident occurs on the shopfloor, especially out of regular office hours.

- **Awareness training:** People remain a major security weakness. Awareness training is necessary both for shopfloor workers and for senior management. Tabletop exercises and simulations can be used to test processes across the organization.

Anomaly Detection
A further method which is demonstrated is continuous monitoring of the OT network for anomalous traffic. An anomaly is an observed behavior which deviates from an established norm and which may indicate a cyber event of interest. Examples include the addition of unexpected devices to the network or unexpected commands at the level of industrial control commands. If a machine is normally started on Monday morning and shut down on Friday evening, then a series of stop commands on a Wednesday afternoon would be an anomaly. Intrusion Detection Systems (IDS) for OT exist to spot such anomalous behavior. However, they require an operator to continuously monitor and evaluate their output. Most producers do not have the skills or the capacity to provide this operator. In the demonstrator we show monitoring as a service: basic operation is performed by a service provider who only informs the responsible production process if a significant event occurs. In the demonstrator the production process is represented by the dashboard.

Anomaly detection addresses current events. It should be complemented by threat intelligence (TI). TI is essentially a form of knowledge sharing across the cyber community to provide users with advance warning of known or emerging threats. Again, the continuous provision and interpretation of TI goes beyond most manufacturers' capabilities. It lends itself to provision as a service-based input to predictive trustworthiness.

Anomaly detection and TI are powerful but in comparison to many safety techniques they have a very high latency. This raises the general question of how safety and security can be holistically addressed in a production environment when they are working to different time scales. Broadly spoken, functional safety works with low latency and low context while cybersecurity works with higher latency and higher context. As described below, cyber alerts can be used as predictive indicators of possible future disruption while safety focuses on immediate, real-time responses.

5. Predictive Trustworthiness

As described above, trustworthiness, consisting of safety, security, privacy, reliability and resilience, needs to be treated holistically in Industry 4.0. Nevertheless, to reach overall trustworthiness for production, different approaches are required in detail at the system, machine, and component levels. At the component level certain critical parts are often protected against cyberattacks, but less critical components are overlooked or the criticality of components is incorrectly assessed. In the case of an incident such as a cyberattack or a failure impacting one component, the question arises how this influences the other components and the whole machine. Currently, the consequence of most incidents is to stop the machine, as is the case with functional safety. However, this behavior may unnecessarily decrease productivity, especially in flexible and modular setups. In this way, unforeseen events are likely as not all permutations of the machine configuration can be assessed during the design phase.

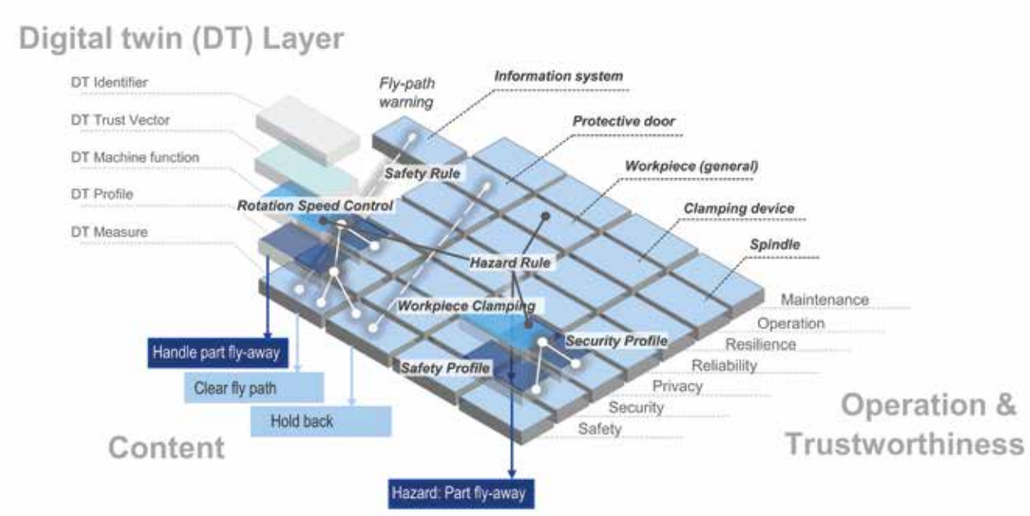

Figure 5: Depiction of the hazard rule part fly-away in a sketch of a digital twin [8]

To improve the situation at machine level, it is first necessary to describe the machine, the components and their functionality in more detail. For example, concerning a lathe, the spindle drive (a component) offers the function of turning the workpiece to the machine level (together with other associated components like a controller and sensors) or the chuck provides the function of clamping the workpiece. As described in [8], this information about the functionality of a component is stored by knowledge graphs in the digital twin of the machine. Furthermore, there are so-called hazard rules at machine level, which associate possible hazards with components and functions. In this example, a workpiece could fly away if the clamping fails. This is captured in a hazard rule, see also Figure 5. A hazard rule also contains calculations for the risk, which depends on the rotation speed and workpiece weight in this case. All hazard rules and safety related aspects of a function are gathered in a so-called safety profile, whereas security related aspects are stored in a security profile.

Given this enhanced knowledge about the machine and the functions, the possible consequences of an incident or a cyberattack can be estimated. Therefore, security related issues, as described in the security profile, can be linked with possibly safety relevant consequences, that are described in the safety profile of the function. This is made possible by the connections of the knowledge graph, which link the information or profiles of a function, including its components, and ultimately enable the entire machine to be linked. Thus, this concept allows the system to estimate the safety relevant hazards for people and the environment, that possibly result from a cybersecurity incident. Hence, this concept fulfills the novel requirements set by many safety standards, that the potential hazards introduced via intentional manipulation or cyberattacks should be assessed and consequently mitigated.

For the mitigation of risks and maintenance of productivity, so-called safety rules, which can react to risks dynamically, are introduced. Safety rules connect hazard rules and safety measures, while dynamically intervening at runtime if necessary. If, for example, the clamping system reports an issue, the calculated risk of the hazard rule “part fly-away” will rise. But instead of just stopping the machine, the corresponding safety rule can calculate a lower but safe speed instead. This is possible on machine level by considering further information from the digital twin like the workpiece weight and the strength of the safety door. Therefore, safety is preserved at runtime while still being productive.

If it is possible to exclude humans from the endangered area of the factory on system level, there are even more possibilities for a reaction to an incident at runtime. Depending on the machine's internal information, the overall system information, but also the severity of the incident, it might be possible for the machine to decide not to slow down or stop at all. This prevents unnecessary

productivity losses due to minor incidents. However, to make a trustworthy decision on the reaction to minor issues, holistic information and a knowledge model are mandatory. Otherwise, stopping or at least slowing down for the sake of safety is necessary.

Based on these methods it is additionally possible to establish predictive trustworthiness. This means, that a change to the machine or a likely incident can already be evaluated in a simulation and so the possible consequences can be predicted. In certain use cases this means that safety critical situations, that would otherwise lead to stopping the machine due to functional safety, are uncovered in simulation and then prevented before they occur. Depending on the simulation capabilities provided by the digital twin, it may already be possible to use this concept during the design phase of the machinery. This configuration avoids costly re-work during construction and commissioning. Another application of this principle is the replacement of a component, where predictive trustworthiness can be helpful to find a fitting and safe replacement part. Another example is machine maintenance, where the question arises as to what safety measures are required in this unusual system state. Additionally, predictive trustworthiness can help security experts to estimate the consequences of possible cyberattacks. Given this information, it can be easier to take appropriate measures to protect the components from the worst possible consequences for safety or the machine. In this case the higher latency of much security information does not impede effective prediction.

6. Implementation Example

At the Hannover Messe 2024, we will demonstrate the interaction of the individual components of trustworthiness in the handling and transportation of dangerous goods with the example of lithium-ion batteries. This use case is based on the SF-KL model process described in [7], in which a model truck is built and assembled from nub bricks. In our use case, we load the truck trailer with dangerous goods (see Figure 6) and show how dangerous goods and their specific transportation conditions can be handled individually and intelligently without affecting the overall production process or safety. We also demonstrate the importance of security in the overall context.

Figure 6: Representative battery pack in one of our truck trailers [7].
(The red batteries represent damaged batteries, the black ones are in good condition)

Our production island _KUBA is a multi-agent system that uses digital twins and the Asset Administration Shell (AAS) to create a digital representation for each physical element. This digital representation enables individual control and monitoring of the production and transportation process of dangerous goods. The first step in the use case is to introduce an empty truck trailer into the _KUBA production island system. The trailer is initially classified as non-hazardous, which enables normal productivity without any special restrictions. As soon as the empty trailer is filled with the lithium-ion batteries in the assembly module, the status in the system changes: the product AAS now detects a potential hazard from the batteries, which requires the trailer to be handled appropriately. All other shuttles are not affected by this.

In order to make the demonstration feasible at a trade fair, we rely on an optical check of the battery packs (see Figure 6) to simulate faults instead of real measurements. In real production these faults can be detected by continuously monitoring the status of the batteries. If a fault is identified, the affected battery pack is classified as actively dangerous. The product AAS is updated to reflect the changed hazard class and immediately initiates safety protocols, such as slowing down the specific shuttle for error analysis and correction by an employee. Once corrected and quality checked again, the product is re-categorized so that the individual speed can be increased again, and the production process can be completed. (Figure 6: Representative battery pack in one of our truck trailers [7] The red batteries represent damaged batteries, the black ones are in good condition.

To ensure the safe handling of dangerous goods, a specialized architecture based on digital twins is essential, but equally critical is ensuring the accuracy and integrity of the data. Safeguarding through cybersecurity methods is essential because without this continuous monitoring, safety can be compromised. Practical examples of potential threats include not only cyberattacks that can compromise the network infrastructure and cyber-physical production systems (CPPS) from the outside, but also internal processes such as maintenance work or the introduction of unknown devices into the system by maintenance engineers – whether unintentionally or intentionally. Especially in the field of dangerous goods transportation, where system-critical parameters form the backbone of safety mechanisms, reliable cybersecurity is of the utmost importance. A scenario in which a fault is detected in the lithium-ion batteries during operation must not lead to the failure of the ejection mechanism due to parameter errors. Such a failure could undermine safety and, in the worst case, lead to serious consequences such as a battery fire, with potentially catastrophic consequences for the production facility. In order to manage the complexity of modern production environments and the associated risk potential, a continuously active anomaly detection system is integrated into our system. This forms a central component of our security architecture and continuously monitors the system parameters and the connection of new and possibly unknown devices to the network. When irregularities are detected, automated warning messages are generated to inform the operator immediately.

Humans play a key role in our safety concept: as the final authority, they make the decisions and initiate the necessary measures based on the information provided by the anomaly detection system. This process enables targeted fault analysis and helps the operating personnel to identify and understand the causes of any problems that occur.

Figure 6 illustrates the architecture of our "Dangerous Good Transportation" Use-Case, which shows the interaction of the digital twins within the multi-agent system. It shows how the AAS provides the necessary information and logic for real-time monitoring and control for each resource and product. Anomaly detection, as part of the orchestrator, plays a central role in this network and works hand in hand with runtime trustworthiness to ensure safe and efficient handling of dangerous goods. Adaptations are made according to the specific hazard classes to respond to any anomalies or changes in system status.

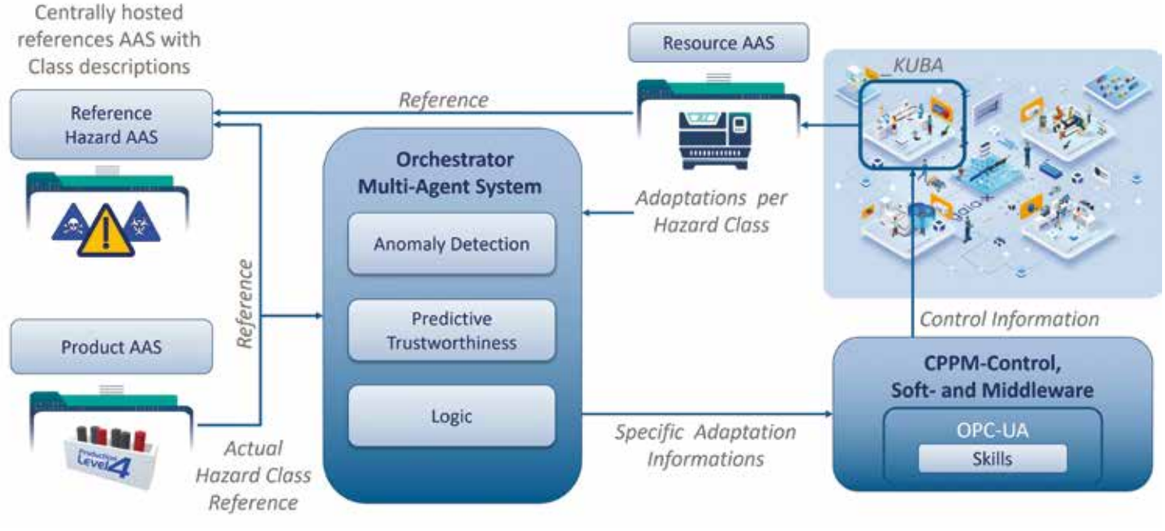

Figure 7: Use-Case dangerous good transportation architecture

7. Conclusion

As shown throughout the paper it is urgently needed to strive for a convergence of safety and security on the shopfloor. This is exactly what the trustworthiness approach achieves. The classic, static approaches to safety no longer meet the challenges and requirements of modern machines and factories, even with the addition of standard off-the-shelf security methods. It was therefore necessary to create a completely new, independent layer of machine trustworthiness with focus on safety and security, which also accesses the information of classic functional safety, but leaves its functionality and properties completely untouched.

To this end, trustworthiness incorporates not only the information from classic functional safety but also other "non-certified" information from the machines, the factory and the entire production environment. Based on this information, trustworthiness can identify emerging risks for people, machines, products and processes at an early stage, even predictively. It means the consequences of incidents or changes at the machine are known and, with the proposed system, even usable for autonomous reaction. Hence appropriate countermeasures are initiated before the intervention of classic functional safety stops the machine. In this way, predictive trustworthiness increases the flexibility, availability, productivity and sustainability of modern machines and factories. Additionally, the trustworthiness approach already takes the requirements of the new European Machinery Regulation [2] into account, which will come into force from 2027 on for new machines to be introduced to the European market.

The next step will be to further maximize the availability of production resources and at the same time enhance the safety system. The aim is to add a safety agent as an additional safety layer that is embedded in a machine and is horizontally connected to other machines or systems. The safety agent will leave functional safety untouched. Plus, a safety agent will be able to listen to messages from other safety agents. This can either be information in the form of lists with planned activities or live information on an action of another safety agent. It will be possible to integrate data from multiple safety agents and from other information systems like Manufacturing Execution System (MES) and Enterprise Resource Planning (ERP) in a common data space. The horizontal and vertical exchange of information between the participating systems will occur via OPC UA and vendor-neutral data collectors.

8. References

[1] Ausschuss für Betriebssicherheit: Cybersicherheit für sicherheitsrelevante Mess-, Steuer- und Regeleinrichtungen. Technische Regel für Betriebssicherheit (TRBS) 1115-1 (2022).

[2] European Parliament: Regulation (EU) 2023/1230 on machinery. (2023).

[3] European Parliament: Directive (EU) 2022/2555 on measures for a high common level of cybersecurity across the Union - NIS 2 Directive. (2022).

[4] National Institute of Standards and Technology (NIST): The NIST Cybersecurity Framework (CSF) 2.0. (2024)

[5] IEC 62443: Security for industrial automation and control systems. (2018).

[6] IAnto Budiardjo, Jon Geater, Frederick Hirsch, Michael Pfeifer, Detlev Richter: Assuring Trustworthiness in Dynamic Systems Using Digital Twins and Trust Vectors. Digital Twin Consortium (DTC) Foundational Paper (2022).

[7] ITechnologie-Initiative ***SmartFactory*** KL e.V.: ***Production Level 4*** - Der Weg zur zukunftssicheren und verlässlichen Produktion. Whitepaper SF-5.1 (2022). Available online: https://smartfactory.de/wp-content/uploads/2022/05/SF_Whitepaper-Production-Level-4_WEB.pdf.

[8] ITechnologie-Initiative ***SmartFactory*** KL e.V.: Das Konzept Knowledge Graph zur Umsetzung von Safety in Digitalen Zwillingen. Whitepaper SF-3.4 (2022). Available online: https://smartfactory.de/wp-content/uploads/2022/05/SF_Whitepaper_SmartSafety-WEB.pdf.

[9] ITechnologie-Initiative ***SmartFactory*** KL e.V.: Safety an modularen Maschinen. Whitepaper SF-3.1 (2018). Available online: https://smartfactory.de/wp-content/uploads/2018/04/SF_WhitePaper_Safety_3-1_DE_XS.pdf.

[10] IT.J. Williams: The Purdue Enterprise Reference Architecture. IFAC Proceedings Volumes, Volume 26, Issue 2, Part 4 (1993).

[11] INational Institute of Standards and Technology (NIST): Guide to Operational Technology (OT) Security. NIST SP 800-82 Rev. 3 (2023).

[12] IISO/IEC 27001: Information security, cybersecurity and privacy protection – Information security management systems – Requirements. (2022).

Versionshistorie
Whitepaper SF-5.1: 05/2022

Herausgegeben von
Technologie-Initiative SmartFactory KL e.V.
Trippstadter Straße 122
67663 Kaiserslautern

T +49 (0)631 20575-3401
F +49 (0)631 20575-3402

Die Technologie-Initiative SmartFactory KL e.V. (*SmartFactory*^{KL}) ist ein gemeinnütziger Verein des öffentlichen Rechts, eingetragen im Vereinsregister Kaiserslautern.

Vereinsregisternummer: VR 2458 Kai

Vorstand
Prof. Dr. Martin Ruskowski (Vorsitzender)
Andreas Huhmann, HARTING AG & Co. KG
Eric Brabänder, Empolis Information Management GmbH
Dr. Detlev Richter, TÜV SÜD AG

Wissenschaftlicher Koordinator
Dr.-Ing. Achim Wagner
T +49 (0)631 20575-5237
M achim.wagner@smartfactory.de

Quellenangabe, Bilder
SmartFactory^{KL} / A.Sell