PLATTFORM INDUSTRIE 4.0

WHITE PAPER

Towards technologies for trustworthiness in value chains

Framework for supporting trust decisions

Imprint

Publisher
Federal Ministry for Economic Affairs and Climate Action (BMWK)
Public Relations
11019 Berlin
www.bmwk.de

Editorial responsibility
Plattform Industrie 4.0
Bülowstraße 78
10783 Berlin

Status
April 2023

This publication is available for download only.

Design
PRpetuum GmbH, 80801 Munich

Picture credit
Nelli Velichko / iStock / title
AndreyPopov / iStock / p. 5
JuSun / iStock / p. 9
gorodenkoff / iStock / p. 12
hh5800 / iStock / p. 16

Central ordering service for publications of the Federal Government:
Email: publikationen@bundesregierung.de
Tel.: +49 30 182722721
Fax: +49 30 18102722721

Table of contents

1 Introduction

Digitalization has undoubtedly become a central factor for transformations in many aspects of personal life and society. While the technical basis enabling the digital transformation, enhanced data collection capabilities, computation and communication resources, is basically the same, digitalization is often addressed within the realms of a domain, such as eGovernment, eHealth, but also "Industrie 4.0" focusing on the digital transformation of the industrial sector. Realizing the need for a broader approach regarding digitalization, the japanese government advocated the concept of Society 5.0. Similar to Industrie 4.0 describing the next evolution of industrial production, Society 5.0 envisions the next evolution of social life and economy. The core technologies for both are very similar, e. g., digital twins, CPSs, 5G networks, cloud/edge computing, IoT, and artificial intelligence. Reaching the goals of both Society 5.0 and Industrie 4.0 requires data centric approaches where cyberspace and physical space tightly integrate and data can be distributed between multiple stakeholders across company and country borders. In such a scenario data security, e. g., confidentiality, integrity and availability, becomes a critical aspect, as it determines whether data exchange actually enhances the quality of products and services or, in the worst case, provides too much of a risk to pursue. However, not only data security must be guaranteed, but also security of the supply chains providing customers with products and services. Even well before the digital transformation products were increasingly manufactured and delivered to end customers by a collaboration of companies. With the digital transformation this development increased rapidly, as it offers various advantages, such as flexibility and cost-effectiveness. In 2022 the JDTF identified four trends regarding the safety and security of supply chains in Society 5.0 [39]:

T1 Digitalization of supply chain and after-sales service: B2B transactions in supply chains are increasingly becoming digital, facilitating the evolution of manufacturing processes, eCommerce, and also after-sales services, such as remote monitoring and maintenance, heavily relying on data provided by cyberphysical systems.

T2 Continuous supply of product and service value in response to diversifying needs and changing external conditions: Supply chains must be organized in an increasingly flexible manner to enable quick reactions to new customer needs, new products, while ensuring the availability of products and services in case of supply chain disruptions, for example, due to natural catastrophes, wars, pandemics.

T3 Changes in structure of supply chain, diversification of industries in which suppliers participate and increasing complexity of supply routes for products and services: Digitalization facilitates the restructuring of supply chains and allows new suppliers to enter the market to compete. This enables companies to diversify their supply routes of core components and services, but also allows suppliers and service providers to become part of supply chains in new domains.

T4 Acceleration of rule formation that requires compliance throughout supply chains: Supply chain participants have diverse set of rules they have to be compliant with. Depending on the type of product or service or the location where said product or service is manufactured/provided at least legislative rules (e. g., GDPR in EU) and domain-specific rules (e. g., automotive, medical) apply. Increasingly, companies want to prove that their products are manufactured in an environment-friendly manner and with sustainability in mind, which has an effect on the whole product/service supply chain.

While these trends have not been formulated explicitly in Industrie 4.0, they are to some degree an implicit part of many typical Industrie 4.0 scenarios. The "Plattform Industrie 4.0" itself explains the term Industrie 4.0 using six example scenarios[1], which all can be mapped to the described trends (Table 1).

In Industrie 4.0, the trust within the entire supply chain can be achieved by utilizing trust-enabling technologies. This white paper presents a framework for supporting trust decisions based on trustworthiness evidence in value creation networks in an industry setting.

[1] https://www.plattform-i40.de/IP/Navigation/EN/Industrie40/WhatIsIndustrie40/what-is-industrie40.html

Table 1: Mapping of "Industrie 4.0" example scenarios to described trends, as conducted by the Plattform Industrie 4.0

	T1	T2	T3	T4
Flexible production		✓	✓	
Convertible factory	✓	✓		
Customer-oriented solutions	✓	✓		
Optimized logistics			✓	✓
Use of data	✓	✓		✓
Resource-efficient circular economy				✓

Secure and safe supply chains are an important part of Industrie 4.0. This white paper proposes a solution for safety and security in digitalized supply chains focusing on the concept of trust and trustworthiness in digitalized environments. In the following chapters an introduction to trustworthiness will be given, followed by the solution architecture and a brief overview over other domains and standardization bodies where trust in supply chains or similar concepts are being discussed.

Trustworthiness in digital supply chains has been a topic of interest in "Plattform Industrie 4.0 for some time merging into two whitepapers in cooperation with the Japanese RRI [1][2] and standardization work in ISO/TC 292 and the corresponding national standardization bodies.. The concepts presented in this document are complementing the previous work by "Plattform Industrie 4.0 and RRI, briefly introduced in 2.2, sharing the nomenclature and basic understanding of trustworthiness and its application in digitalized supply chains. While the previous work focused on protocols and concepts to establish trustworthiness, for example, in e-procurement, and the complex multiparty trustworthiness relations along supply chains in a top-down approach, the work in this whitepaper provides a bottom-up approach, showing how trust-enabling technologies (3) can be utilized to support trustworthiness in supply chains.

The document at hand is targeted at the Industrie 4.0 community with a focus on developers and standardizers working on trustworthiness solution and standards, as it provides an overview over current developments in the standardization bodies on the topic of trustworthiness and trust-enabling technologies. Finally, we introduce a technology-driven architecture (Supply Chain and Trust Management Architecture), based on trusted computing, to acquire, process and distribute trustworthiness-related data.

2 Trust and Trust Decisions in Value-Creation Networks

2.1 Concept of Trust and Trustworthiness

Stock and Boyer [3] study several definitions of supply chain management to arrive at what they call, *a consensus definition:*

> The management of a network of relationships within a firm and between interdependent organizations and business units consisting of material suppliers, purchasing, production facilities, logistics, marketing, and related systems that facilitate the forward and reverse flow of materials, services, finances and information from the original producer to final customer with the benefits of adding value, maximizing profitability through efficiencies, and achieving customer satisfaction. [3]

With modern supply chains in information-driven societies, many of the entities increasingly involve artificial intelligence [4], [5], which Boden describes as "computers that do the sorts of things that minds can do" [6]. Although the role of trust in societies, involving humans, animals and inanimate entities including artificial intelligence, may not be immediately discernible, Marsh et al., [7] refers to the work by Luhmann [8], which suggests that trust helps reduce the complexity of everyday life by allowing us to take certain things as given.

In their 2022 work on trust and trustworthiness in artificial intelligence [9], Lewis and Marsh classify the definitions of trust found in existing literature into four perspectives: (1) the accountability perspective; (2) the coordination perspective; (3) the goodwill perspective; and (4) the decision perspective.

Bryson [10] concludes that "we need to know we can hold the human beings behind that system to account". Ryan [11] goes as far as stating that complex machines should not be viewed as trustworthy! Lewis and Marsh, disagreeing with this notion that there is no validity of trust when dealing with non-humans, extrapolate the *accountability perspective* as an additional requirement that a trustee be able to morally accept blame. This, the authors admit though, necessitates a theoretical reasoning different from the well-known *decision perspective* of trust, which we describe later.

The *coordination perspective* sees trust as a coordination of expectations and that it only exists in a relationship between peers. In this respect, Bryson [10] defines trust as "a relationship between peers in which the trusting party, while not knowing for certain what the trusted party will do, believes any promises being made". Baier [12] assumes cognitive and linguistic ability in both the trustor and the trustee in order to establish a process of expectation reconciliation. Sako and Helper [13] term "contractual trust" between participants as an expectation to keep to an established agreement between the parties. Reina [14] uses the term "transactional trust", where trust is only given in order to develop or maintain trustworthiness. Lewis and Marsh find this implied human exceptionalism of the coordination perspective between peers limiting and "incompatible with a functionalist view on intelligence and agency" [9]. For example, this perspective does not even extend to trust between humans and animals, let alone between humans and machines.

The *goodwill perspective* can also be found in the works of Baier [12], Sako and Helper [13], Lagerspetz [15] as well as the more recent work from Sutrop [16] in which trust is seen to exist when the trustor's reliance depends on the goodwill of the trustee. According to this perspective, the trustor may feel that the trustee is capable of meeting the former's needs through the latter's actions; and as a result the trustor approaches the trustee with trust. As Dennett [17] found, this intentional viewpoint – that an entity is able, willing or even designed to meet another's needs – is particularly useful when reasoning about complex objects.

Lastly, the *decision perspective* is based on the most widely accepted notion of trust [18]–[21] wherein trust is seen as an "act of choosing to put oneself into a situation of risk, where the outcomes are dependent on the actions of another" [9]; and this act of trusting may be so done based on a set of evidences. In the decision perspective, trust is observed to be highly contextual, idiosyncratic and subjective [22]. Dwyer [22] argues that the act of trusting should be empowered (i. e., given as a choice), not enforced (i.e., told to do so). This is particularly relevant to us in the context of supply chains where either humans or machines (or both) are responsible to trust (as an action) other humans or machines. The trustor should be able to make the trust decision based on the trustor's own policies (which may encapsulate a level of subjectivity), instead of having a trust decision enforced due to certain metrics or attributes that the potential trustee presents. Besides trustors being able to apply their own policies to make a trust decision, we also explore how some of that decision making can be informed by the system (the trustee; specifically a *trusted system* according to RFC 4949 [23]) that "operates as expected, according to design and policy, doing what is required – despite environmental disruption, human user and operator errors, and attacks by hostile parties – and not doing other things". It is to be noted that evidences about this *reliable system* operating "as expected" are subject to their completeness (with respect to conformance policies) as well as freshness (with respect to reflecting reality accurately at a given point in time).

Following the definition of trust from the *decision perspective*, trustworthiness is an important input that goes into a trust decision [21], [24]. Trustworthiness is an attribute of the trustee, which guides whether a trustor chooses to trust the trustee. Trustworthiness "is based on any number of traits, beliefs, desires, intentions, competencies, and so on" [9]. Potential trustors form their subjective beliefs about the trustworthiness of potential trustees based on the evidence about the latter, something McKnight and Chervany [25] called "trusting beliefs". It is to be noted that regardless of a trustee's trustworthiness, a potential trustor may choose to trust the trustee anyway.

Trustworthiness is defined in the ISO/IEC 20924:2021 standard, as the "ability to meet stakeholder expectations in a demonstrable, verifiable and measurable way". This relates to the concept of "contractual trust" [13] for its focus on expectations but the ISO/IEC 20924:2021 departs from the human exceptionalism of the coordination perspective, described above, by stating that trustworthiness, as an attribute, can be applied to humans and their groups but also to entities that are abstract and non-human, such as services, products, data and information. This ability [of the trustee] to meet [the trustor's expectations] in demonstrable, verifiable and measurable ways can be derived or evaluated from: (1) observations of how the trustor behaves in the relevant context; as well as (2) insights, e. g. through transparency, into how the trustee works, thinks and values, with respect to the expectations of the trustor. Lewis and Marsh [9] term the former is called black-box and the latter white-box evidences of trustworthiness that serves as an input to trust as a decision. In our recent work on trust in manufacturing [26], we present both of these evidences by describing manufacturing as a process model (i. e. insights of what is being done) as well as evidences of the actual execution of the process (i. e. observations).

A technical basis in support of the *ability [of the trustee] to meet [the trustor's expectations]* in the Internet is produced by the RATS WG in the IETF. The working group's charter states that:

> In network protocol exchanges, it is often the case that one entity (a relying party) requires evidence about the remote peer (and system components [RFC4949] thereof), in order to assess the trustworthiness of the peer. (IETF RATS WG [23], [27])

By enabling devices or systems (trustees) to take on the role of an attester they can produce digital evidence about their trustworthiness characteristics. Evidence production of attesters is based on well-known roots of trust included in the attester. Corresponding trustor entities rely on two complementary roles: (1) the burden of appraisal of evidence, which can be rather voluminous and device or system specific, is taken on by the verifier role; and (2) the relying party role can then act on the attestation results produced by trusted Verifiers based on policies that apply to it. Attestation results are typically more concise and universal and therefore easier to process and act upon by various types of relying parties. In general, trust in a role is established via well-known trust anchors. Consecutively, external entities, such as a verifier, can establish and manage trust relationships with roots of trusts via their corresponding trust anchors.

An attester produces white-box evidence. Exposing such evidence to a verifier requires a trust relationship between attester and verifiers. Analogously, the relying party that requires a trustworthiness assessment about remote peers requires trust relationships between itself and the verifier. Additional trust relationships extend into the supply chain entities that manufactured the attester (trustee): a verifier requires a trustworthy description of an attester's intended conformance characteristics and composition, which means there must be policies in place that establish trust relationships between external endorsers of the attester or corresponding reference value providers.

In essence, remote attestation is not a trivial process and requires various actors to collaborate. Remote attestation can be a vital part of the assessment of *trustworthy characteristics* of a trustee and requires multiple *trust decisions* about which entities to trust in the automation of the corresponding procedures handled by trustors. In order to automate remote attestation, it is also vital to take into account that automation of trustworthiness assessments depends on trust decisions manifesting in policies that can support such automation. To enable the implementation of global and scalable systems, the number of manual trust decisions must remain at a manageable level. Additionally, trust decisions and their basis must be made transparent, auditable (in a given scope), and – in a best case scenario – are standardized and interoperable.

2.2 Trustworthiness in Supply Chains

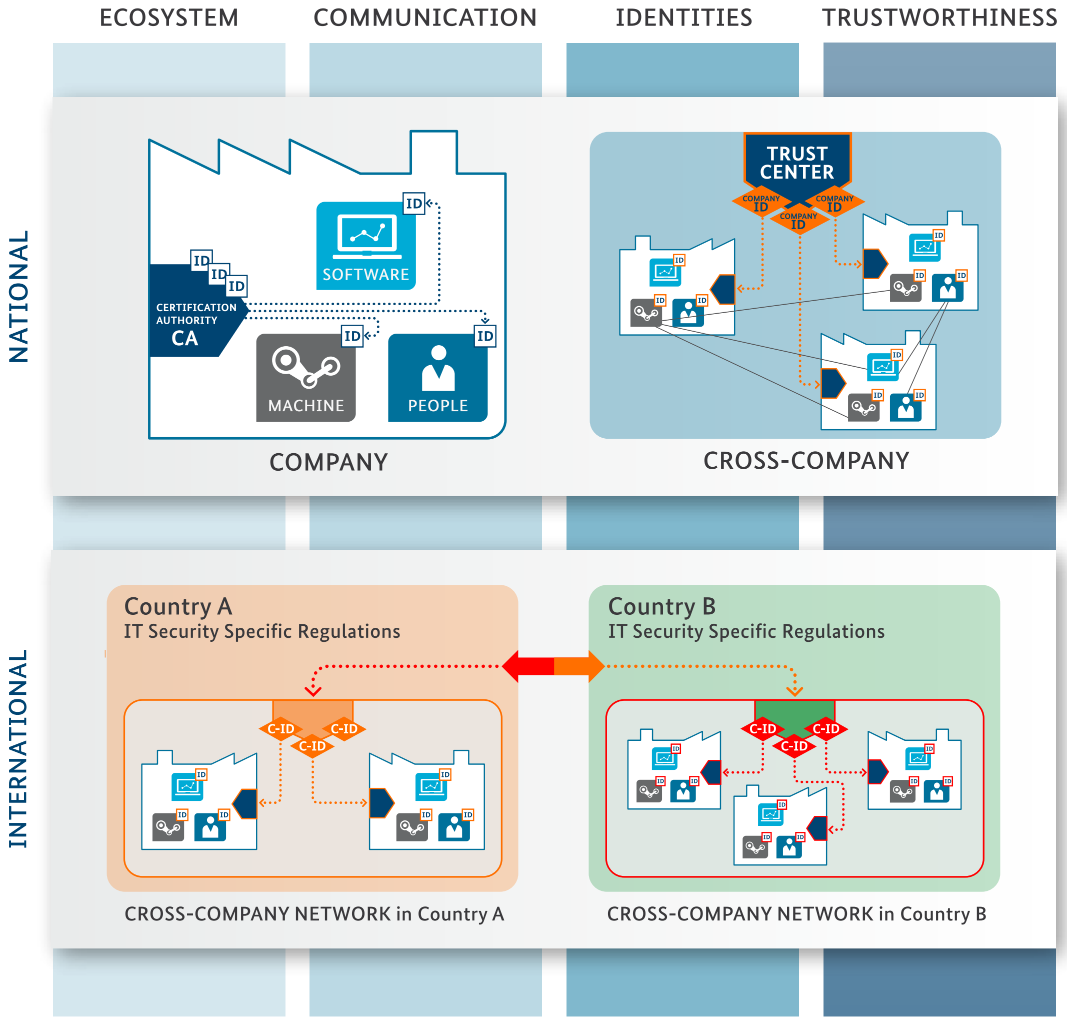
The topic of trustworthiness in supply chains of "Industrie 4.0" has been discussed previously in [1] and [2]. Both publications give an introduction to the general concept of trustworthiness and its importance for supply chains, relevant international standards and propose solutions to certain aspects of trust in supply chains.

The joint whitepaper between RRI (Japan) and "Plattform Industrie 4.0" (Germany) focusses on IIoT value chain security and the role of trustworthiness [1]. The authors concentrate on the possibility of creating ad hoc trustworthy relationships between companies, who did not have a joint business history before. They aim to provide support to companies, so that they can find trustworthy collaboration partners and establish truthful relationships. The overall vision is presented in Figure 1. The authors identify mechanisms to support trustworthiness assurance, propose the TECEP, and the idea to utilize trustworthiness profiles.

The publication was followed by a document with a focus on chain of trust for organizations and production [2]. A supply chain typically consists of many entities, who contribute to the product available for the end user. This means that there are also many potential attack vectors. The aim of the publication was to provide support to manufacturers, so that they can find trustworthy components easily and can establish ad hoc trustworthy relationships with supply chain partners. The analysis led to a structured approach to achieve trustworthiness, a comprehensive understanding of the chain of trust topology and general requirements for a chain of trust. The document showed especially the role of globally unique IDs, digital signatures and a digital proof of process to determinate the trustworthiness of organizations and products.

Both publications and the document at hand share a common basis and understanding for the trustworthiness in supply chains. As such, the work presented in this document shall be interpreted as a continuation of the shared efforts to enable and facilitate trustworthiness mechanisms in modern supply chains. Similar to the preceding work, this document introduces a solution to another aspect of trust in supply chains, which has not yet been addressed.

Figure 1: Overall scenario of Plattform Industrie 4.0 production [1]



Supply chains today are (almost) fully automated, realized with computer systems and devices that control, manage, and operate supply chain processes. If we want to trust supply chains, we hypothesize that we have to trust the digital systems involved. So, the overall question becomes: How can we trust a device?

Digging deeper, this yields the question: Which “parts” of a device can we put trust in at all? The answer is quite simple: hardware and software. We have to have confidence (trust) in all the hardware and all the software involved in every part of the supply chain. This includes all involved supply chains for all parts.

This leads to two further questions and answers:

1. How can we trust hardware? We need to make sure (or know for sure) that the device – or a piece of hardware – was manufactured and shipped as expected.

2. How can we trust software? We need to make sure (or know for sure) that the software was developed and is executed as expected.

“As expected” means that all items and steps in the supply chain can be *trusted*, i.e., they must provide verifiable evidence of its defined (and successfully completed) “treatment” process (production, shipping, etc.).

3 Trust-Enabling Technologies

There exist several technologies to enable trust in computing systems. In this section, Trusted Computing (3.1) and Confidential Computing (3.2) are introduced.

3.1 Trusted Computing

Trusted Computing enables systems to produce verifiable evidence about all its running software, including boot components and OS software. Trusted Computing based on a hardware root-of-trust has been developed by the industry to protect computing infrastructure and billions of endpoints. The TCG has developed the TPM with cryptographic features that enforce certain behaviors and protect systems from unauthorized modification, malware, rootkits, and other attacks.

Standards-based Trusted Computing technologies developed by TCG members now are deployed in enterprise systems, storage systems, networks, embedded systems, and mobile devices, and they can secure cloud computing and virtualized systems. As a result, systems, networks, and applications are not only more secure, less susceptible to viruses and malware, and more reliable, but also easier to deploy and manage.

Several definitions of trust in different domains have been evaluated in 2. The definition of trust from the TCG is:

> "An entity can be trusted if it always behaves in the expected manner for the intended purpose."
> (TCG 2004)

Chris Mitchell defines Trusted Computing as:

> "With Trusted Computing, the computer will consistently behave in expected ways, and those behaviors will be enforced by computer hardware and software." (Chris Mitchell [28])

3.1.1 Trusted Platform Module (TPM)

The TPM [29] is a cryptographic coprocessor in hardware with secure storage and secure key generation capabilities, and it is hardened against physical attack. A TPM is a passive device, and software must actively use it. The BIOS or UEFI, the bootloader, and other components must have (built-in) code to use and control the TPM. The Linux kernel and Windows have TPM drivers on board to communicate with the TPM.

A TPM can be leveraged for the protection of credentials, such as cryptographic keys; EA policies enable for the fine-grained definition of access policies. TPM remote attestation is used for the *detection* of compromise to a system. That is, a TPM and its ecosystem enable a system to provide (explicitly) verifiable evidence about all loaded (software) components to third parties. The TPM extends the RTM, and implements an RTS as well as an RTR.

3.1.2 Measurement and Attestation Roots (MARS)

The TCG further specifies MARS that maintains a reliable and remotely verifiable device identity, rooted in hardware. The boot health – the integrity of the components used to boot the device – is also remotely verifiable. Like the TPM, MARS provides a RTM, a RTS, and a RTR. In contrast, it lacks enhanced features of a TPM, such as secure sessions and EA policies. MARS is a specification that is not tied to any particular hardware. It rather constitutes a concept and a protection profile that must be implemented in the target hardware. So, its application is hardware-specific. MARS enables explicit remote attestation, like the TPM and its use cases are primarily targeted at constrained IoT systems. MARS enables explicit remote attestation, like TPM, and its use cases are primarily focused on constrained IoT systems.

3.1.3 Device Identifier Composition Engine (DICE)

DICE is another standard from the TCG and targets very constrained devices where a TPM or MARS cannot be used for technical or economic reasons. It provides a strong hardware identity as well as firmware integrity. The DICE engine creates a compound device identifier from a unique hardware ID and a hash of the first loaded firmware (stage 0 firmware). DICE provides a RTR and supports implicit remote attestation. Like MARS, DICE is a specification that is not tied to any particular hardware. It must be implemented in the target hardware, making its application hardware-specific.

3.2 Confidential Computing

Confidential computing enables new public cloud scenarios (e.g., migrating extremely sensitive data to the cloud, and enabling multi-party sharing scenarios that have been difficult to build due to privacy, security, and regulatory requirements). [30]

Today, data is often encrypted at rest, in storage, and in transit across the network, but not while in use in memory. Additionally, the ability to protect data and code while it is in use is limited in conventional computing infrastructure. Confidential Computing utilizes TEEs to achieve code confidentiality, authenticated launch, and protected/isolated execution of user-defined code.

> The Confidential Computing Consortium is a community focused on projects securing data in use and accelerating the adoption of confidential computing through open collaboration. [...]
>
> The CCC brings together hardware vendors, cloud providers, and software developers to accelerate the adoption of Trusted Execution Environment (TEE) technologies and standards. [31]

Confidential computing has complementary goals with Trusted Computing that provides load-time integrity verification of all software components of an entire system. However, Trusted Computing does not protect data while in use in memory.

3.2.1 Trusted Execution Environments (TEEs)

A TEE is an environment for securely executing code. Code running in a TEE can have high levels of trust. Programs that are executed inside a TEE can ignore threats from the rest of the system. TEE programs are called TAs. The counterpart to a TEE is the REE, such as Linux, Microsoft Windows, or Apple MacOS.

Today's systems tend to run a lot of software. And the more software is executed, the larger the attack surface becomes and the higher the probability of security vulnerabilities. To work around that, the TEE concept was developed. TEEs are specified by a consortium of concerned parties: network operators, manufacturers, OS vendors.

There are many interpretations of what is meant by trust in a TEE. In the TEE it is used to imply that you may have a higher level of trust in validity, isolation, and access control on items (assets) stored in this space. This leads to the following assertion: "Trusted OS and TAs executed inside that space are more trustworthy."

On the one hand there are manufacturer and use case specific TEEs. They are designed to restrict themselves to a limited use case, hereby meeting the needs of a specific manufacturer. On the other hand there are "real" TEEs. They have high-quality internal isolation and are designed to enable a device's best security to be leveraged by developers beyond the initial production-line installers.

3.2.2 Arm TrustZone

Arm TrustZone is system-wide approach to embedded security. It is available in Arm Cortex-based processor systems. Cortex-based cores are used in everything from microcontrollers (MCUs) to high-performance CPU. TrustZone is an embedded security technology that starts at the hardware level. It creates two environments that can run simultaneously on a single core:

- a "Secure World", and
- a not-as-secure world ("Normal World")

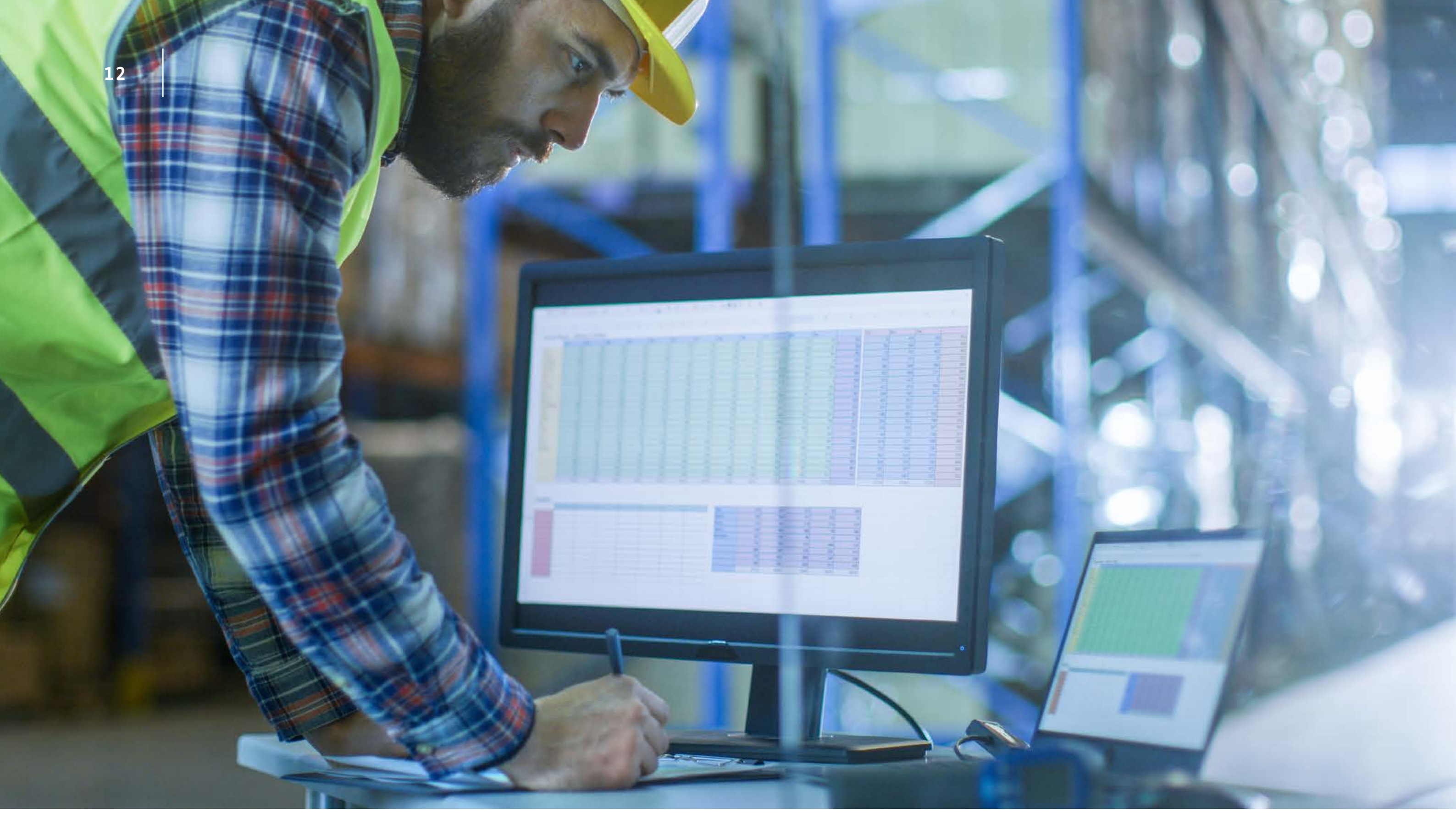
Increasingly, developers need to secure systems beginning at the lowest levels, i. e., the physical layer, which includes the boot process. A TrustZone-enabled system starts the bootloader of the normal world first after the secure world has fully booted. Two virtual processors execute code in a time-sliced fashion. Context switching is realized through a new core mode (monitor mode) when changing the currently running processor. The software in monitor mode is implementation defined, and saves and restores states when switching worlds.

3.2.3 Intel Software Guard Extensions (SGX)

Protection rings as implemented by a CPU protect the OS from applications, and applications from one another. However, applications are not protected from privileged code attacks. A malicious application may exploit a flaw in the OS kernel and gain access through privileged code. Goals of Intel SGX are:

- An application defends its own secrets.
- Keep the attack surface small (application parts + CPU).
- Require no separate processor.

Intel SGX is an extension to CPU instructions introduced with Intel Skylake processors. Its key concept is the "enclave", a protected environment that contains the code and data of security-sensitive computations. A system can run multiple enclaves; there is no limitation. While Arm TrustZone has exactly one trusted world, Intel SGX has multiple. Enclaves are isolated from the untrusted software outside – including the OS – as well as from other enclaves.



4 Utilizing Trust-Enabling Technologies

Integrity is a crucial aspect in today's distribution and supply chains. Ensuring the integrity of supply chains means to ensure the integrity of any involved computer system. That is, it must be ensured hardware and software is authentic and integral.

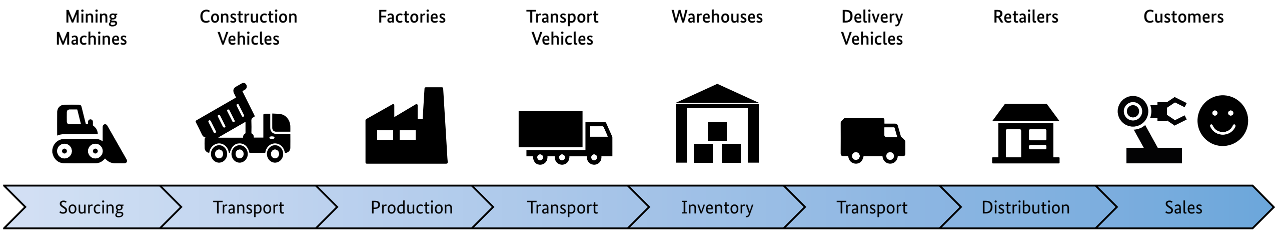
Trust-enabling technologies are already available in many hardware and software systems in production, OT, and IT but are not always being utilized as state of the art.

4.1 introduces the supply chain model as used in this white paper. In 4.2 the adapted VCP is described that contains trustworthiness criteria. 4.3 introduces our Supply Chain and Trust Management architecture that augments supply chains with trust aspects and includes VCPs.

4.1 Supply Chain Model

Figure 2 depicts an idealistic supply chain for an industrial robot. In (today's) supply chains, many parties are involved in each of the steps, all having their own supply chains, resulting in nested supply chains.

Figure 2: Simplified supply chain model



4.1.1 Hardware Supply Chain

This section examines the hardware supply chain by first looking at how trust is currently established, and then identifying problems in the form of attacks.

4.1.1.1 Trusting Hardware – State of the Art

The device manufacturer makes sure that the manufacturing process is "secure". Usually, this process is not under the control of an external party, e. g., the customer. The device manufacturer ensures that all installed components are genuine and issues a platform certificate that confirms that the platform is assembled according to the data sheet. The platform certificate is made available to the customer. A platform certificate is an attribute certificate, not an identity certificate, so no public key is included. The platform certificate is typically stored in the persistent storage of the TPM. The TPM manufacturer issues an EK Cert and stores it in the TPM. The EK is a unique hardware key in the TPM; the private portion never leaves the TPM.

The device manufacturer may "seal" devices' housings by chemical or mechanical means, such as a white paint that "bleeds" red when cut or scratched. Or the manufacturer may use tape or seals that show evidence of removal. Further, frangible (brittle, breakable) covers or seals are other methods available using current technology.

In summary, it is essential to trust the manufacturers to do their job properly, because it is not possible to gain insight into their processes and the machines and software used.

4.1.1.2 Hardware Supply Chain Attacks

There is a huge number of (potential) supply chain attacks. This may be compromised production equipment that is infiltrated by a hardware "trojan", or software manipulation in the assembling machine. PITM attacks are omnipresent, even for hardware. On its way to the customer, the product/device is opened, chips or software are replaced with malicious ones, or secrets are extracted.

4.1.2 Software Supply Chain

This section examines the software supply chain by first looking at how trust is currently established, and then identifying problems in the form of attacks.

4.1.2.1 Trusting Software – State of the Art

Software vendors *should* ensure a robust and secure software development process. Like it is the case with hardware, (proprietary) software is developed "behind closed doors". Hence, the customer or any other third party has no access to the software's source code. The software development and testing process is entirely up to the software vendor. The software may even be (semi-)formally designed, verified, and tested. Testing the software (fuzzing, unit and integration tests, etc.) is also a common practice. Some vendors sign their software binaries in order to ensure authenticity and integrity, making the certificate available to the customer, including necessary intermediate and root certificates. The installation on-premise is sometimes restricted to only signed software to be installed.

Like with hardware, it is essential to trust the manufacturers to do their job properly, because it is not possible to gain insight into their processes and the machines and software used.

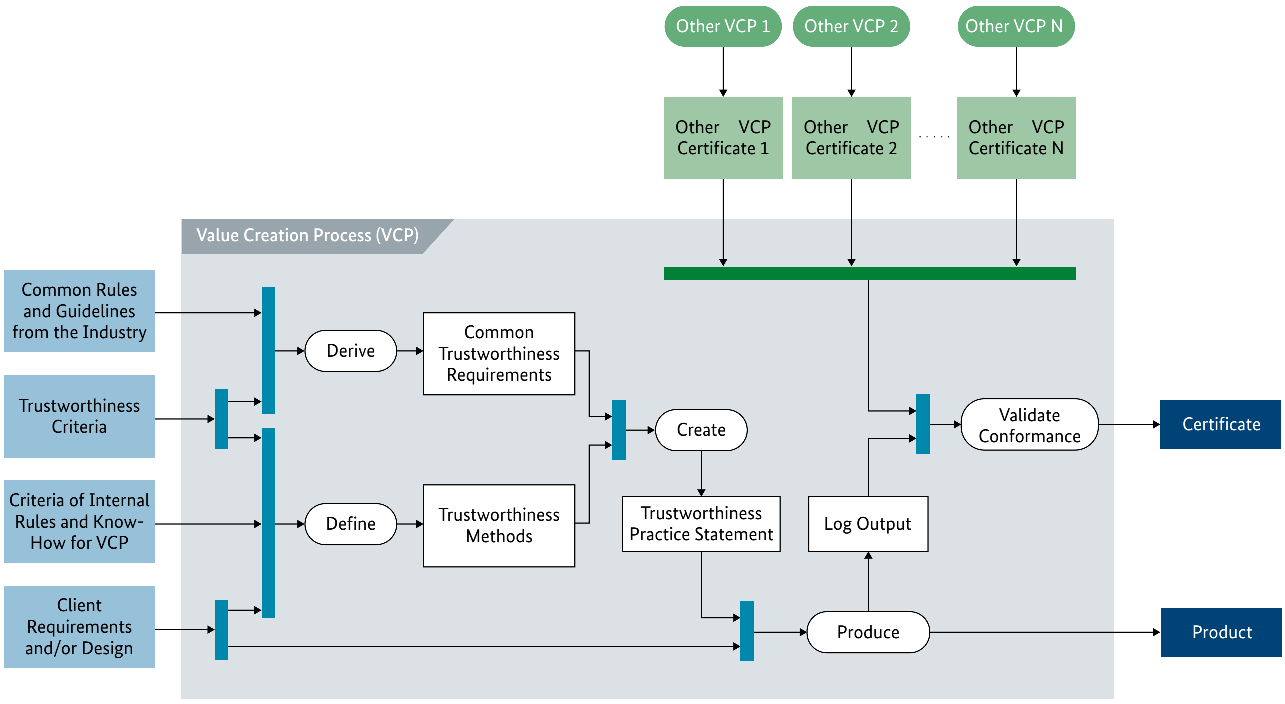
4.1.2.2 Software Supply Chain Attacks

There is countless (potential) attacks on software. These may be compromised compilers and build tools, or even the software vendor's build machine be corrupt, compiling a virus into the binary. Code signing could also be corrupt, resulting in signing a malicious binary.

PITM attacks can take place on several layers and locations. The vendor's root certificate may be replaced with a malicious one, or the binary may be replaced with a malicious one during transmission.

There is also data-at-rest attacks where an attacker replaces a binary while the machine is powered off. The attacker may place a malicious root certificate in the trusted certificate store. Runtime attacks may be exploited, due to a bug in the software binary that, e. g., causes privilege escalation or arbitrary code execution.

Figure 3: Value Creation Process (VCP)


Other VCP 1
Other VCP 2
Other VCP N
Other VCP Certificate 1
Other VCP Certificate 2
Other VCP Certificate N
Value Creation Process (VCP)
Common Rules and Guidelines from the Industry
Trustworthiness Criteria
Criteria of Internal Rules and Know-How for VCP
Client Requirements and/or Design
Derive
Common Trustworthiness Requirements
Define
Trustworthiness Methods
Create
Trustworthiness Practice Statement
Log Output
Produce
Validate Conformance
Certificate
Product

4.2 Value Creation Process

A generic VCP is depicted in Figure 3, which is based on the work by Nagayoshi et al. [26]. It is focused on transparency and verifiable evidence of the production process, which leads to trustworthiness of the process. It takes as input common rules and guidelines from the industry, Trustworthiness Criteria, criteria of internal rules and know-how for the VCP, and client requirements and/or design. From these, common trustworthiness requirements are derived, and trustworthiness methods are defined. Next, trustworthiness practice statements are created from these. The final product is produced from the client requirements and/or design and one or more trustworthiness practice statements. Production-related data is collected during this process to serve as evidence for the production process and as input to the conformance validation. Conformance validation uses the data from the production process, as well as certificates from other VCPs – including those from other vendors, such as installed microchips – to create a certificate for the manufactured product.

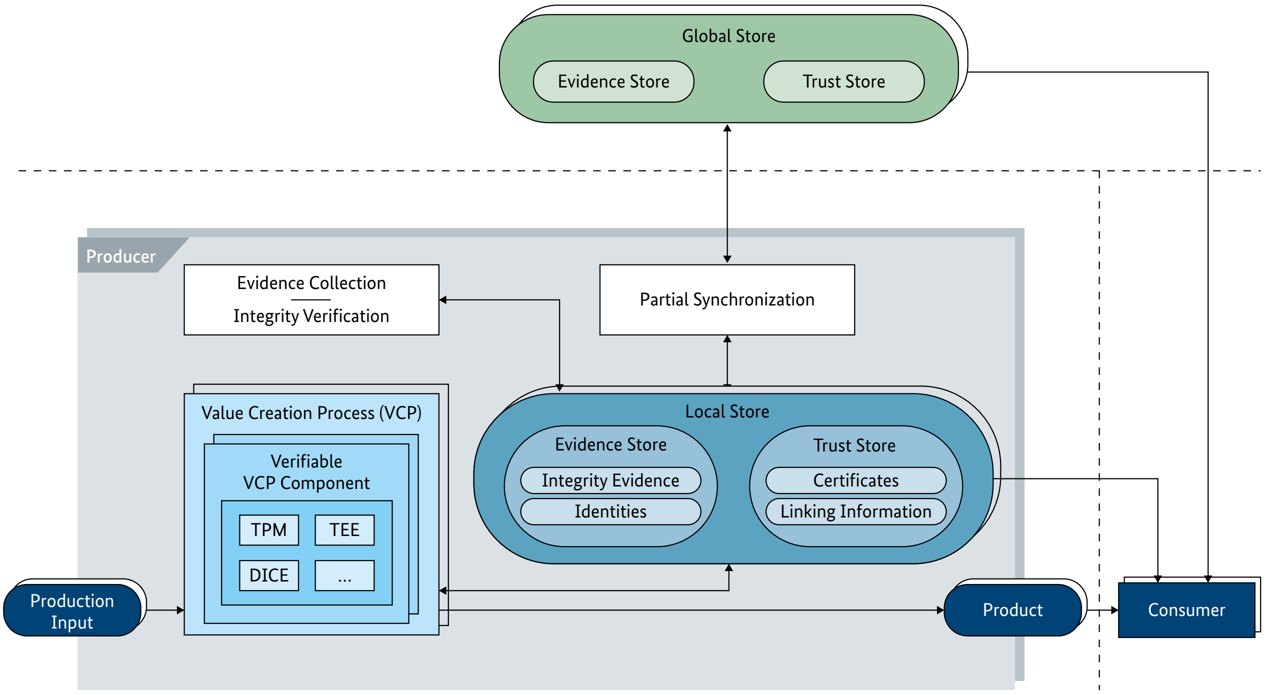
The output of the manufacturing process as depicted in Figure 3 is the product and the certificate. Note that these can also be multiple certificates and products, depending on the particular VCP.

4.3 Supply Chain and Trust Management Architecture

The goal of this section is to present a generic supply chain architecture that is complemented by trust measures: the Supply Chain and Trust Management architecture (Figure 4).

The central element is a company that takes production input, such as silicon, microchips, housings, or software libraries, and produces a product by going through several VCP components. One or more products are then delivered to one or more consumers. The consumers can be end customers (B2C), other companies (B2B), or transport companies.

Figure 4: Supply Chain and Trust Management Architecture

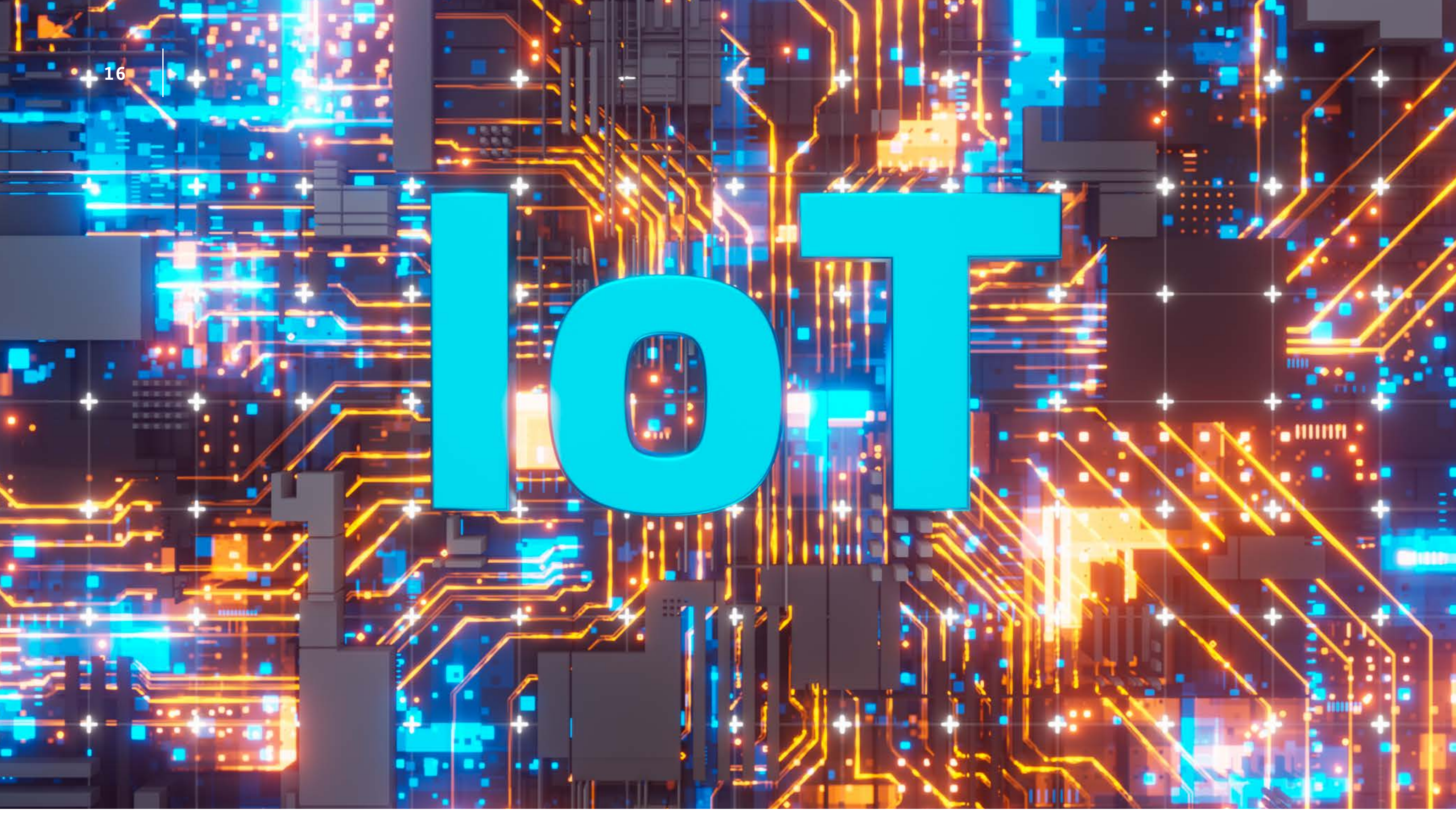


The production process (Value Creation Process VCP) consists of several VCP components. Such a component can be any machine, software, or cyberphysical (production) system, that is being utilized during production. To make the VCP component verifiable, it needs to have any form of trust-enabling technology (section 3) available. The architecture proposed is hereby technology-agnostic and supports approaches like TPM, TEE and others to be as inclusive as possible with regards to these technologies. Remote attestation, as described in section 3, plays an integral role verifying the production systems integrity and hence, the trustworthiness of generated data. The evidence collection and integrity verification collects integrity data from all involved VCP components, records all this data, and stores it in the local evidence store for later processing, such as certificate generation and auditing.

The local trust store contains certificates from the production process as described in 4.2. This includes linking information to assembled parts from other manufacturers. For example, a robot manufacturer certainly has certificates from all purchased parts from another manufacturer in its trust store, so that the certificate for the final product – the robot – contains links to the certificates of involved parts of the robot (e. g., display unit, power supply, or motors).

Certificates from other manufacturers may be obtained from global evidence and trust stores, or may be part of the production input. Produced certificates and evidence of rather confidential nature may be shared with consumers directly, and not put into a global store. However, non-critical evidence data and certificates may be synchronized (partially) with global stores.

All evidence and certificate stores may be realized as common (graph) databases. Another option would be the usage of Distributed Ledger Technologies (DLT), such as blockchain, and a consensus protocol – e. g., the Practical Byzantine Fault Tolerance (pBFT) – to realize a secure distributed database. As transparency and integrity are the main goals for certificate and evidence stores, DLTs are viable candidates for the use case.



5 Related Work

This section lists related work in the context supply chain security, trust, and trustworthiness.

5.1 Asset Administration Shell

Asset Administration Shell (AAS) [32] [33]is a core concept developed by the Plattform Industrie 4.0. AAS is an implementation of digital twins for IIoT, providing a standard for virtual models reflecting physical objects by linking this virtual model to the asset. The AAS is the standardized digital representation of the asset and provides capabilities for asset interoperability. Technical functionality that is exposed by an asset can be described in the AAS, as well as asset-related data. This is organized in standardized submodels for various aspects of the asset. The concept of submodels aims to define submodels for any concern, for example submodels for safety and security. Regarding the proposed architecture in Section 4, a standardized submodel for trustworthiness would provide a suitable integration in a AAS and therefore facilitate the integration of trustworthiness in other use cases for IIoT. The Trustworthiness Profile, as defined in [1], would be part of the trustworthiness submodel of an asset, signaling trustworthiness expectations and capabilities along the supply chain.

5.2 Generic Trust Anchor API

The concepts introduced in the document at hand heavily rely on the usage of trust-enabling technologies as introduced in section 3. In order to facilitate and spread trust and trustworthiness in supply chains, it is advisable to not limit the concept technologically by focusing on only a limited set of vendors or trust-enabling technologies. While all those technologies have their unique features, there is also a large common set of capabilities. Currently, a developer of industrial applications must implement security functionality which may be provided by secure elements in a technology-specific manner using custom programming interfaces and libraries for each trust-enabling technology. To facilitate and spread the usage of trust-enabling technologies, ISO/IEC JTC1 SC41 is specifying TS 30168 “Generic Trust Anchor Application Programming Interface for Industrial Applications. The goal is to simplify the integration and usage of secure elements in IIoT devices by standardizing an Industrial IoT Security API. This will be achieved by providing a technology-agnostic API set for commonly used cryptographic functionality of trust-enabling technologies, such as random numbers, encryption, decryption, and signatures. The work organized under ISO/IEC JTC 1/SC 41 is carried out by national standardization bodies such as DIN in Germany.

5.3 Framework for establishing trustworthy supply chains

ISO TC 292 WG4 is currently working on a guideline for a "Framework for establishing trustworthy supply chains" (AWI 22373)[2]. It defines an approach to support stakeholders in a supply chain to accomplish a chain of trustworthiness regarding properties of identifiable material goods along the supply chain. It gives guidance on the identification of trust domains, their corresponding trustworthiness attributes, and define a standardized data structure to exchange trustworthiness relevant information for signaling assurance to trustworthiness properties, between different supply chain nodes. It will support typical trustworthiness relevant properties, such as interoperability, robustness, accountability, transparency while preserving privacy, etc.

5.4 Supply Chain Integrity, Transparency, and Trust (SCITT)

IETF is a standards defining organization that produces standard building blocks for the Internet. Currently, a new WG forms in the IETF: the SCITT WG. A summary of the mission statement is:

> SCITT is an initiative to define industry standards on how to provide interoperable, concise, air-gap'able trust assertions in support of believable accountability and auditability. (IETF SCITT WG [34])

The charter is still under development at the time of writing this white paper and can be found at the IETF website[3] [34]. The development is documented in an IETF email list (scitt@ietf.org) and in a GitHub repository[4].

In more detail, the SCITT WG defines two essential message types that enable global, uniform signing mechanisms for authentic and accountable statements about supply chain artifacts (Figure 5). The first message type is called a *Claim* and wraps opaque statement payloads in a standardized, interoperable, and state-of-the-art signing envelope. The second message type is called a *Transparent Claim* that is composed of a Claim and a corresponding countersigning proof about that Claim being notarized by a Transparency Service, called a Receipt. Notarization involves two essential activities:

- Storing the Claim in an append-only log – typically a Merkle tree – and
- creating the countersignature that includes a Merkle Inclusion Proof in a Receipt.

In summary, Receipts and Claims in composition constitute Transparent Claims. Hashes of all issued Claims stored by the Notary (the actor conducting the notarization activity) are added to a Merkle tree. A Merkle tree operated on by a Notary is called a Registry.

Both signed statements and countersigned Merkle Inclusion Proofs are based on the IETF COSE [35]. The serialization of COSE is based on the CBOR [36]. The use of COSE and CBOR enable applications involving constrained devices in constrained node environments – that constitute the majority of the IoT – as CBOR and COSE are lightweight, require small processing power, minimal stack sizes, and have a small data-in-flight footprint. Analogously, Receipts are very small, can be stored independent of their corresponding Claims and, if trusted (e. g., by a well-managed trust anchor store), can be validated offline in air-gap usage scenarios. Auditability of Receipts or Transparent Claims is enabled by Transparency Services who implement the role of a Notary and maintain a corresponding Registry.

The flexible and scalable trust relationships of SCITT are tied to trust decisions about Transparency Services. If a supply chain entity trusts a Transparency Services, on the one hand it simply can validate a receipt even without having the capability of processing the statement about supply chain artifacts, e. g., itself. On the other hand, a supply chain entity can present the receipt to a relying party (colloquially, a consumer of SCITT Transparent Claims or Receipts), which then can conduct a full audit trail of related statements coming from the same Issuer via Transparency Services. As no statements can ever be removed from a Registry, and statements never expire, extensive audit trails can be facilitated by SCITT Transparency Services.

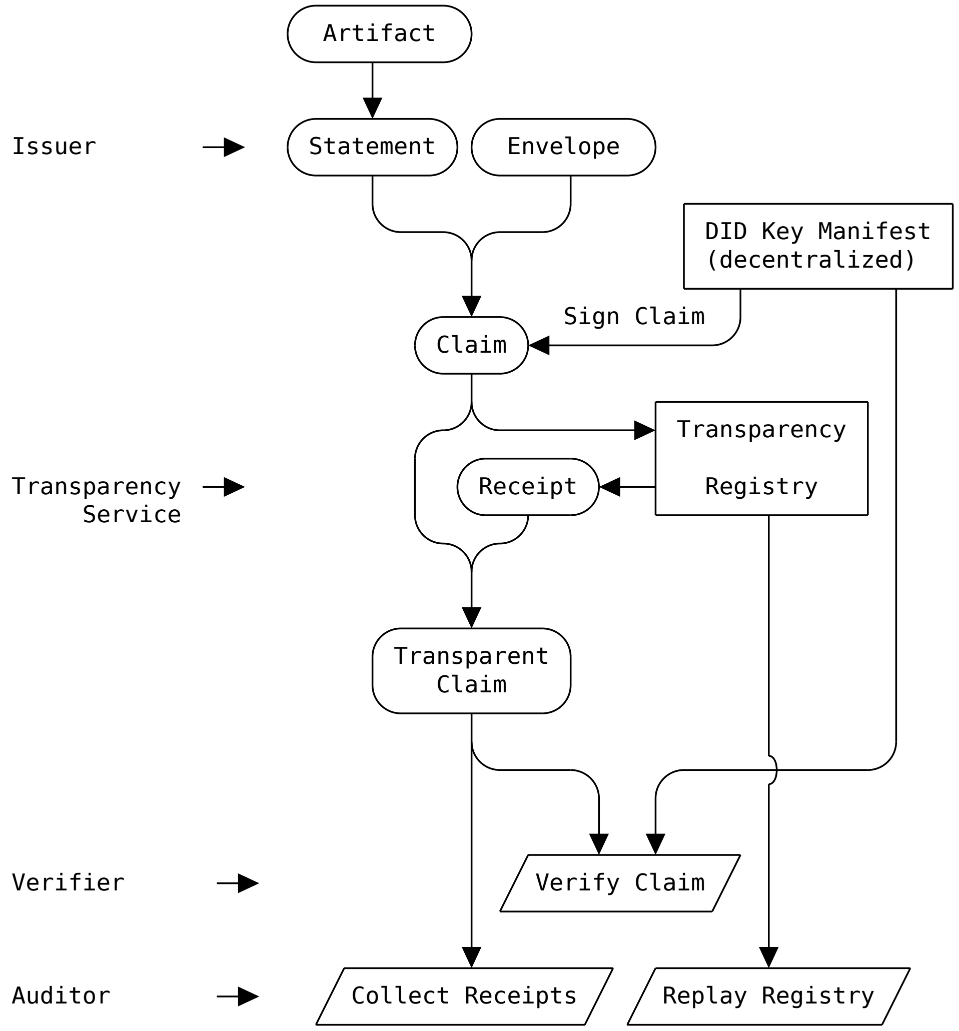
[2] https://www.isotc292online.org/

[3] https://datatracker.ietf.org/wg/scitt/about/

[4] https://github.com/ietf-scitt/charter

The first use case to be addressed by the SCITT WG is the Software Supply Chain. While the SCITT payload is opaque to the mechanisms of the Transparency Services, a well-known set of software supply chain documents exists – called software bill of materials (SBOM) – that will drive the first use case of the WG.

Figure 5: IETF SCITT architecture (work in progress) [34]



5.5 Software Bill of Materials (SBOM)

On May 12, 2021 US president Biden issued EO 14028 on

> Improving the Nation's Cybersecurity [37]

In Section 4 *(Enhancing Software Supply Chain Security)* a deadline for publication of a *minimum elements for a SBOM* document by NIST 60 days after the release of the executive order was set. This deadline follows the guidance of *providing a purchaser a Software Bill of Materials (SBOM) for each product directly or by publishing it on a public website* included in the executive order.

Ultimately, various deliverables defined in the executive order inform the FAR Council and other appropriate US agencies on how to handle and procure software and how to contract IT and OT services in the future. Accordingly, NIST published on July 12th, 2021 the document *The Minimum Elements For a SBOM*. The NIST document states:

> An SBOM provides those who produce, purchase, and operate software with information that enhances their understanding of the supply chain, which enables multiple benefits, most notably the potential to track known and newly emerged vulnerabilities and risk.
>
> SBOM will not solve all software security problems but will form a foundational data layer on which further security tools, practices, and assurances can be built. [38]

There are three element types that compose the minimal elements for SBOM defined by the NIST publication:

Data fields that provide baseline information about each component that should be tracked, namely: Supplier, Component Name, Version of the Component, Other Unique Identifiers, Dependency Relationship, Author of SBOM Data, and Timestamps,

Automation support, including automatic generation and machine-readability to allow for scaling across the software ecosystem, and

Practices and processes that define the operations of SBOM requests, generation and use including Frequency, Depth, Known Unknowns, Distribution and Delivery, Access Control, and Accommodation of Mistakes.

There are also three SBOM formats acknowledged by the NIST publication in support of automation: SPDX, CycloneDX, and SWID tags. All formats have strengths and weaknesses today. There is a trend of convergence, though, as, for example, SPDX started to include reference to other formats, such as SWID tags and CoSWID tags as well as *gitbom*. SBOMs are often referred to as *ingredient lists for software*, but their today's complexity and corresponding capability already exceeds the use of a list of characterized subcomponents. Existing SBOM formats also start to cover more software life-cycle phases than the *Release* phase that is targeted by executive order 14048 where a purchaser of a released software has to be provided with a SBOM.

The guidance in the NIST publication on minimal SBOM elements deliberately excludes requirements on integrity and authenticity mechanism that would render SBOM documents believable and trustworthy inputs to supply chain trust mechanisms, such as SCITT. Nevertheless, SBOM documents are one of the first well-defined payloads of SCITT in the context of its first use case: Software Supply Chain.

6 Summary

This white paper presents the blueprint of a supply chain architecture that introduces “trustworthiness” measures into every step of a value chain. Our work focuses on the trustworthiness in the context of “Industrie 4.0”, driven by the “Plattform Industrie 4.0”, and safety and security of supply chains in “Society 5.0” as described by the JDTF. This covers the entire software and hardware life cycle, an important focus topic of the “Plattform Industrie 4.0”.

We introduce the overall topic in section 1 and present trends for “Society 5.0” and their relation to “Industrie 4.0”. In section 2, we take a deeper look at the topics of trust and trustworthiness and go into various definitions. Section 2.2 presents concepts, considerations, and existing work related to trust in supply chains. It highlights key challenges in establishing trust in hardware and software.

In 3, trust-enabling technologies are presented – Trusted Computing and Confidential Computing – both of which play an essential role in establishing trust into hardware and software components in every part of a digitalized supply chain.

A possible architecture is introduced in section 4.3. It describes the supply chain model, the trustworthiness-enhanced VCP, and the Supply Chain and Trust Management Architecture. The Supply Chain and Trust Management Architecture demonstrates how to create and verify trustworthiness evidence as well as the distribution of this verifiable evidence to other participants of a supply chain. Considerations for improving the life cycle of hardware and software are also presented.

Eventually, we compare this work with existing work in 5, such as IETF SCITT [34] and RATS [27] as well as SBOMs.

Literature

[1] A. Kitamura, V. Bellinghausen, J. Fujita, A. Furukawa, L. Jänicke, M. Jochem, W. Klasen, A. Maftun, T. Matsumoto, M. Shiba, N. Suzuki, T. Walloschke, S. Zimmermann, T. Yamada, and T. Yoneda: "IIoT value chain security – the role of trustworthiness", 2020. [Online]. Available: https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/IIoT_Value_Chain_Security.pdf.

[2] V. Bellinghausen, J. Fujita, A. Furukawa, L. Jänicke, M. Jochem, A. Kitamura, W. Klasen, A. Maftun, K. Mahara, T. Matsumoto, K. Rannenberg, M. Shiba, N. Suzuki, and T. Yoneda: "IIoT value chain – security chain of trust for organizations and products", 2022. [Online]. Available: https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/IIoT_Value_Chain_Security2.pdf.

[3] J. R. Stock and S. L. Boyer: "Developing a consensus definition of supply chain management: A qualitative study", *International Journal of Physical Distribution & Logistics Management*, 2009.

[4] C. Dirican: "The impacts of robotics, artificial intelligence on business and economics", *Procedia-Social and Behavioral Sciences*, vol. 195, pp. 564-573, 2015.

[5] V. Kumar, D. Ramachandran, and B. Kumar: "Influence of new-age technologies on marketing: A research agenda", *Journal of Business Research*, vol. 125, pp. 864-877, 2021.

[6] M. A. Boden: *AI: Its nature and future.* Oxford University Press, 2016.

[7] S. Marsh, T. Atele-Williams, A. Basu, N. Dwyer, P. R. Lewis, H. Miller-Bakewell, and J. Pitt: "Thinking about trust: People, process, and place", *Patterns,* vol. 1, no. 3, p. 100039, 2020.

[8] N. Luhmann: *Trust and power.* John Wiley & Sons, 1979.

[9] P. R. Lewis and S. Marsh: "What is it like to trust a rock? A functionalist perspective on trust and trustworthiness in artificial intelligence", *Cognitive Systems Research*, vol. 72, pp. 33-49, 2022.

[10] J. Bryson: "AI & global governance: No one should trust AI", *United Nations Centre for Policy Research*, vol. 21, 2018.

[11] M. Ryan: "In AI we trust: Ethics, artificial intelligence, and reliability", *Science and Engineering Ethics*, vol. 26, no. 5, pp. 2749-2767, 2020.

[12] A. Baier: "Trust and antitrust", *Ethics*, vol. 96, no. 2, pp. 231-260, 1986.

[13] M. Sako and S. Helper: "Determinants of trust in supplier relations: Evidence from the automotive industry in japan and the united states", *Journal of Economic Behavior & Organization*, vol. 34, no. 3, pp. 387-417, 1998.

[14] D. S. Reina: *Trust and betrayal in the workplace: Building effective relationships in your organization.* Berret-Koehler Publishers, 1999.

[15] O. Lagerspetz: *Trust: The tacit demand*, vol. 1. Springer Science & Business Media, 1998.

[16] M. Sutrop: "Should we trust artificial intelligence?" *Trames: A Journal of the Humanities and Social Sciences*, vol. 23, no. 4, pp. 499-522, 2019.

[17] D. C. Dennett: “Intentional systems”, *The Journal of Philosophy*, vol. 68, no. 4, pp. 87-106, 1971.

[18] D. Gambetta: “Can we trust trust?” in *Trust: Making and breaking cooperating relations*, Basil Blackwell, Ltd., 1988, pp. 213-237.

[19] N. Luhmann: “Familiarity, confidence, trust: Problems and alternatives”, in *Trust: Making and breaking cooperating relations*, Basil Blackwell, Ltd., 1988, pp. 94-107.

[20] S. P. Marsh: “Formalising trust as a computational concept”, PhD thesis, University of Stirling, 1994.

[21] H. D. McKnight and N. L. Chervany: “Trust and distrust definitions: One bite at a time”, in *Trust in cyber-societies*, Springer, 2001, pp. 27-54.

[22] N. Dwyer, A. Basu, and S. Marsh: “Reflections on measuring the trust empowerment potential of a digital environment”, in *IFIP international conference on trust management*, 2013, pp. 127-135.

[23] R. W. Shirey: “Internet Security Glossary, Version 2”. RFC 4949; RFC Editor, Aug-2007.

[24] R. C. Mayer, J. H. Davis, and F. D. Schoorman: “An integrative model of organizational trust”, *Academy of management review*, vol. 20, no. 3, pp. 709-734, 1995.

[25] H. D. McKnight and N. L. Chervany, “What is trust? A conceptual analysis and an interdisciplinary model”, in *Proceedings of the americas conference on information systems*, 2000, pp. 827-833.

[26] H. Nagayoshi, T. Terada, T. Sato, A. Basu, C. Otahara, H. Uchiyama, and M. Sato: “Verification of process suitability for creating trust in supply chains”, in *Symposium on cryptography and information security*, 2020.

[27] Internet Engineering Task Force: “Remote ATtestation ProcedureS (rats)”. Internet Engineering Task Force; Internet Engineering Task Force, 2022.

[28] C. Mitchell and I. of Electrical Engineers: *Trusted computing*. Institution of Engineering; Technology, 2005.

[29] Trusted Computing Group: *Trusted Platform Module Library – Part 1: Architecture*, Family 2.0, Level 00, Revision 01.38. Trusted Computing Group, 2016.

[30] E. Lear and H. Tschofenig: “Defining and Enabling Confidential Computing”. Confidential Computing Consortium; Confidential Computing Consortium, Aug-2018.

[31] Confidential Computing Consortium: “Confidential Computing Consortium”, 2022. [Online]. Available: https://confidentialcomputing.io/.

[32] P. I. 4.0: “Details of the asset administration shell, part 1”, 2022. [Online]. Available: https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.pdf.

[33] P. I. 4.0: “Details of the asset administration shell, part 2”, 2021. [Online]. Available: https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part_2_V1.pdf.

[34] Internet Engineering Task Force: "Supply Chain Integrity, Transparency, and Trust (scitt)". Internet Engineering Task Force; Internet Engineering Task Force, Jun-2022.

[35] J. Schaad: "CBOR Object Signing and Encryption (COSE)", RFC Editor; Internet Requests for Comments; RFC Editor, {RFC} 8152, Jul. 2017.

[36] C. Bormann and P. Hoffman: "Concise Binary Object Representation (CBOR)", RFC Editor; Internet Requests for Comments; RFC Editor, {RFC} 8949, Dec. 2020.

[37] Executive Office of the President: "Improving the nation's cybersecurity", 17-May-2021. [Online]. Available: https://www.federalregister.gov/executive-order/14028.

[38] The United States Department of Commerce: "The minimum elements for a software bill of materials (SBOM)", 12-Jul-2021. [Online]. Available: https://www.ntia.doc.gov/files/ntia/publications/sbom_minimum_elements_report.pdf.

[39] Japan Digital Trust Forum: Supply Chain Trust Issues and Recommendations. [Online]. Available: https://jdtf.or.jp/en/report/whitepaper/

LIST OF PARTICIPANTS

Basu, Anirban (Hitachi Ltd.) | Birkholz, Henk (Fraunhofer SIT) | de Meer, Jan (HTW Berlin) | Đukanović, Siniša (Fraunhofer SIT) | Eckel, Michael (Fraunhofer SIT) | Flubacher, Björn (BSI) | Houdeau, Detlef (Infineon) | Jänicke, Lutz (PHOENIX CONTACT GmbH & Co. KG) | Jochem, Michael (Robert Bosch) | Klasen, Wolfgang (Siemens) | Maftun, Aliza (Siemens) | Lieske, Matthias (Hitachi Europe GmbH) | Rannenberg, Kai (Goethe University Frankfurt) | Simo-Fhom, Hervais Clemence (Fraunhofer SIT)

This publication is a result of the working group "Security of Networked Systems" of Plattform Industrie 4.0

www.plattform-i40.de